

Distrust in genetically altered foods

Attempts in Europe to resist imports of genetically modified maize from the United States risk a damaging trade battle. But opposition springs from science and public distrust, both of which need a considered response from industry and politicians alike.

PUBLIC attitudes to the safety of genetically engineered products in general, and foods in particular, are not 'rational' in the strictly scientific sense. Indeed, it would be surprising if they were. The rapid growth of interest in organic foods is only one consequence of the high environmental — and perhaps health — price that has been paid for past enthusiasm for chemical herbicides, insecticides and fertilizers. Yet while many consumers are voicing increasing demands for 'natural' products, farmers are seeking to cut costs and increase yields by moving in precisely the opposite direction through the use of crops that have been genetically tailored for greater efficiency of production.

Conflict has been inevitable. It should therefore come as little surprise that the critics of genetic engineering have seized a new opportunity — the first exports of genetically-engineered agricultural products from the United States to Europe at the end of the current growing season — to highlight their concerns (see page 564). Ironically their protests, timed to coincide with this week's World Food Day, partly reflect the success of the US agricultural industry in persuading regulators that their products are safe to grow. But they have also spotlighted both residual concerns about potential hazards, and the cultural challenge of handling low-level risk on both sides of the Atlantic.

Neither issue is straightforward, although the first is easier to address than the second. Critics of the new crops have raised a wide variety of concerns. Most are already being carefully watched for, such as the legitimate fear that genes for herbicide resistance might pass from a crop such as oil seed rape to its 'weedy' relatives (see *Nature* 380, 31; 1996). Others, such as the claim that stimulating the resistance of crops to chemical herbicides encourages the excessive use of such herbicides, concern broader questions of environmental policy that cannot be tackled through the regulation of genetically-engineered crops alone.

Separate from these is a more specific concern that has surfaced in a particularly acute form over one particular crop. The crop in question is a variety of herbicide-resistant corn (or maize) that has been developed by the Swiss company Ciba-Geigy to express an additional trait, namely toxicity to a major pest, the European corn-borer. The concern, highlighted earlier this year by Britain's Advisory Committee on Novel Foods and Processes (ACNFP), is that a gene resistant to the antibiotic ampicillin, used as a marker in an early stage of the development process, could — at least theoretically — be passed to man via bacteria lodged in the gut of animals which eat the maize unprocessed.

Ciba-Geigy's response to this concern, which has contributed directly to a regulatory impasse in Brussels, is that it is exaggerated. Strictly speaking, the company may be right; certainly the series of events that would need to occur — including the transfer of the offending stretches of DNA to gut bacteria — have a low probability. But the risk, nevertheless, is there, and is a matter of genuine scientific debate. There is certainly wisdom in those members of the ACNFP who argue against attempts to shrug off any potential increase in antibiotic resistance in the population, even if small. European science advisers now addressing this issue should ensure that it receives full consideration in their recommendations.

Handling risks of this type has now become the most difficult

task of regulators on both sides of the Atlantic. Enhancing scientific understanding is essential, but is not the whole solution. Just as challenging, but just as necessary, is the creation of trust. It is that which European consumers currently appear to lack, combining deep-rooted cultural fears of genetic manipulation with past experience of the aggressiveness of some agri-business companies (a tradition which Ciba-Geigy is perpetuating, by all accounts). If both technical and political monitoring procedures can be put in place that are sufficient to generate and maintain this trust, there is no reason why genetically-engineered foodstuffs should not enjoy success in the market. But without such procedures — or even recognition for their need — seed companies and their allies risk growing opposition. □

Spreading assessment skill

The systematic judgement of research performance is a growing industry in search of international quality assurance.

TIGHTENING budgets are putting research institutions under increasing pressure to justify decisions such as who is promoted, which grant application is approved, and which departments should be favoured or stripped of funding. Such institutions also acknowledge a need to evaluate completed research projects.

The value of quantitative measures, such as publication-based indicators and research grant statistics, in both processes is self evident. As was clear at an international meeting last week (see page 567), the way in which they are used can be highly contentious. But the heat of the debates at the conference would have been reduced if delegates, mostly research scientists, had been able to consider the practical experience of different countries.

An analysis of that experience, and of related work conducted by science policy researchers, is urgently needed. There are plenty of potential users. In Europe, research agencies in Italy, Spain and the new democracies are prominent among those trying to work out how best to allocate funds. Their scientists are often to be found engaged in heated discussions about assessment in which the wheel is repeatedly reinvented.

Enter, possibly, the European Science Foundation (ESF). Funded by national research agencies, the foundation is well placed to pull together information and analysis, and to promote a system of 'best practice' in research assessment throughout the continent. Two years ago, after much soul-searching, the ESF decided that science policy should become its core mission. Since then, it has identified research assessment as an important area on which its advice could be of value.

So far, however, even though it possesses the necessary infrastructure and contacts, only good intentions have emerged from the ESF's office in Strasbourg. The time has come for the organization to show more vigour by sinking its teeth into the research assessment problem. In doing so, it would provide a much-needed service not only for Europe, but for the many other countries in which objective research assessment is increasingly considered as an essential goal. □



Fullerenes finally score as Nobel committee honours chemists

London. The Nobel committee last week gave the answer to a favourite topic of speculation at chemistry conferences for several years: when would the chemistry prize be awarded for the discovery that carbon atoms can assemble into the C_{60} carbon cages of buckminsterfullerene — and who would win it?

Both questions have now been answered, the second with the decision to award the prize to Sir Harry Kroto of the UK University of Sussex (bottom right), and Robert Curl and Richard Smalley of Rice University in Houston, Texas (right).

The discovery of buckminsterfullerene, named after the American architect Buckminster Fuller, who used similar geodesic domes to construct large-scale structures, was made at Rice University in 1985 during experiments to investigate carbon clusters formed by laser ablation of graphite.

That the discovery of C_{60} would one day earn Nobel recognition has been widely regarded as inevitable. It is hard to identify any other recent finding in chemistry that has stimulated such intense and diverse activity throughout the physical sciences.

Organic chemists have literally found a new dimension for synthesis in C_{60} 's spherical topology. Solid-state physicists have scaled new heights in molecular superconductivity. Films of solid C_{60} doped with alkali metals are superconducting at up to 33 degrees kelvin, 20 K above the previous record for a molecular superconductor.

C_{60} exhibits novel tribological effects and promising nonlinear optical properties, and is a candidate for the carrier of unexplained features in astronomical spectra. "It has completely changed our perspective on carbon chemistry," says Kroto.

The question was never so much whether C_{60} would win a Nobel prize as when, and — most of all — for whom. David Jones, a chemist at the University of Nottingham, says he had expected that the award might be given "only when C_{60} turned out to be of some use".

But commercial applications remain remote, and Jones considers that C_{60} and the related carbon-cage clusters collectively called fullerenes have so far done little of practical value "other than providing entertainment". (Jones's speculations about a curved form of graphite in his Daedalus column in *New Scientist* in 1966 are commonly

regarded as the conceptual beginning of fullerene science.)

In contrast, Robert Haddon, one of the team at Bell Laboratories in New Jersey who discovered superconducting C_{60} compounds in 1991, feels that the prize "could have been awarded by the end of 1991, when it was clear that fullerenes would change organic chemistry and materials science".

The question of 'who' was the hardest, as the route from discovery to worldwide impact has involved many significant contributions. But no award would have made sense that omitted Kroto, Smalley or Curl.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Robert Curl (above left), his colleague Richard Smalley (on screen behind model), and Harry Kroto (below right) featured on *Nature's* front page (left) in 1985 with their report of a discovery that has transformed knowledge of nanoscale carbon structures.

Kroto's interest in long-chain carbon molecules, called polyynes, as the source of features in the microwave spectra of interstellar clouds stimulated the experiments at Rice, where Curl introduced Kroto to the laser-ablation apparatus that he and Smalley were using to study semiconductor clusters.

The experiments on graphite were initiated to look at clusters of up to around 33 atoms, which were expected to be linear chain molecules. C_{60} and the other cage-like fullerenes were an unexpected by-product. The peculiar stability of C_{60} posed a puzzle that demanded explanation.

But perfecting the conditions required for optimal C_{60} formation was a task performed by graduate students Jim Heath and Sean O'Brien (and also initially Yuan Liu and Qing-ling Zhang) at Rice University. Heath and O'Brien co-authored with Kroto, Curl and Smalley the 1985 paper that reported the discovery (see *Nature* **318**, 162; 1985).

C_{60} might well have stayed the curiosity it remained for the next five years if physicists Wolfgang Kratschmer, Don Huffman and their respective students Kosta Fostiropoulos and Lowell Lamb had not found a way to make the compound in gram quantities in 1990. It was this discovery that allowed fullerene science to blossom.

There is no question that fullerenes have provided entertainment for scientists and non-scientists alike. But how important are they? Kroto feels that the Nobel committee have taken "a gamble that fullerenes will be one of the biggest things in the 21st century". In Haddon's view, the significance is in fundamental, not applied, science. "It's the kind of fundamental breakthrough in chemistry of the kind we don't really see any more," he says.

Curl is confident that technological uses will come in the next few decades. But he suggests that the most telling factor now is that fullerenes illustrate "the interplay of [molecular] topology and structure". Fullerene science has led to the discovery of a host of materials, carbon-based and otherwise, whose atomic-scale structure dictates particular nanoscale shapes and forms.

Carbon nanotubes, discovered in 1991, are the most celebrated example: tubes of graphite-like carbon a few nanometres in diameter whose closed end caps are curved by the same five-membered-ring defects that induce the curvature of the C_{60} cage.

Smalley agrees that "the utility of C_{60} is in what it has taught us about nanoscale carbon structures" — namely that the graphite sheet "is a pretty cool thing". Fullerenes may turn out to be not so much an end in themselves as the beginning of a new view of materials science at the interface of the molecular and the bulk scales.

Philip Ball

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Geoff Tomkinson/SPL

Physics award acclaims superfluid helium

London. The 1996 Nobel prize for physics has been awarded to Douglas Osheroff of Stanford University, California, and to David Lee and Robert Richardson of Cornell University, New York. Their 1971 discovery of superfluidity in the helium isotope ^3He spurred two decades of research into the behaviour of macroscopic systems at very low temperatures, at which the effects of quantum mechanics are most evident.

Superfluids are unlike normal fluids. Coffee stirred in a cup eventually comes to rest, its energy dissipated by internal friction. But a superfluid has no internal friction. Consequently, it flows without dissipation, slips easily through cracks too small for normal liquids to penetrate and, under some conditions, can even flow uphill.

Superfluids derive these strange properties from a quantum mechanical effect known as Bose-Einstein condensation, in which a large proportion of the fluid's particles all begin to move in an identical fashion.

Many elements would in principle exhibit superfluidity, but freeze solid before they can be cooled to a sufficiently low temperature. Helium is unique in remaining liquid even as its temperature approaches absolute zero. It is therefore the only element in which the effect occurs. Superfluidity was first discovered in 1938 in liquid ^4He , the most abundant isotope of helium on Earth.

For many years it was believed that the much rarer isotope ^3He would not behave similarly. According to quantum mechanics, particles fall into two classes, 'fermions' and 'bosons', depending on the value of an

liquid of fermions, any small attractive force between particles would bind them into boson-like pairs, which then could undergo Bose-Einstein condensation. This theory explained the low-temperature behaviour of many metals and alloys, in which electrons pair and lead to superconductivity — a close analogue to superfluidity. Its success suggested that superfluidity might occur in ^3He , the atoms of which attract one another weakly by van der Waals forces.

"There were many theoretical predictions of a superfluid transition in ^3He ," says Tony Leggett, a theoretical physicist at the University of Illinois, pointing out that some estimates of the transition temperature were as low as 10^{-17} K. But, despite much effort, even by 1970 superfluidity had not been detected in ^3He .

In 1971, Osheroff was Lee's graduate student, and Richardson was a young faculty member at Cornell. The three were investigating nuclear antiferromagnetism in solid ^3He and, according to Lee, made their discovery "more or less by accident".

The group's work required a magnet weighing several tons — a rare commodity among researchers. In autumn 1971, when another group at Cornell needed the magnet, Osheroff decided to test a new refrigeration device that the group had acquired. Helium-3 is unusual in that energy is needed to change it from liquid to solid, a characteristic that can be exploited to cool a mixture of liquid and solid helium to exceedingly low temperatures.

The group devised an experiment in which a solid-liquid mixture was cooled at a constant rate. The experiment went as planned until they saw an unexpected fluctuation in the cooling rate when the temperature reached 2.6 mK. Not thinking about superfluidity at the time, the team published a paper suggesting that the effect might

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

David M. Lee (left) and Robert C. Richardson, both of Cornell University, toast their Nobel success.

reflect a phase transition in the organization of spins in the solid.

They then carried out a series of nuclear magnetic resonance experiments to isolate the effects of the solid and liquid in the system, and found that several phase transitions were occurring. One was solely in the liquid. Eventually the group began to suspect that they were seeing a Bardeen-Cooper-Schrieffer transition from a normal fluid to a superfluid.

Osheroff says that he still has his laboratory notebook in which the data were recorded. Like many graduate students, he often worked in the middle of the night. An entry at 2.40am on 20 April 1972 reads: "Have discovered the BCS transition in liquid ^3He tonight."

Ironically, the paper reporting these results was initially rejected by *Physical Review Letters*. One referee argued that the system "cannot do what the authors are suggesting it does". But the authors appealed, and the paper was eventually published.

Further experiments by other groups revealed that ^3He was far more complex than expected. "Once the experiments started coming out, it soon became clear that theorists had failed to calculate all the really interesting things," says Leggett. For example, ^3He has three distinct superfluid phases, while ^4He has only one.

Mark Buchanan

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

AP/Paul Sakuma

Douglas Osheroff with apparatus that he uses to study properties of liquid helium.

attribute called 'spin'. Whereas ^4He atoms are bosons, ^3He atoms are fermions. And, because only bosons can undergo Bose-Einstein condensation, superfluidity in ^3He appeared unlikely.

But in 1957 John Bardeen, Leon Cooper and J Robert Schrieffer showed that, in a

Japanese pioneer recalls 'prescient paper'

Tokyo. Japanese chemist Eiji Osawa was devastated when he saw the paper published in *Nature* in 1985 (see previous page) describing the discovery of buckminsterfullerene, and comparing its structure to that of a soccer ball.

Fifteen years previously, when at Kyoto University, Osawa had predicted the possible existence of a C_{60} molecule with a structure like a soccer ball. But he gave up his work — which was published only in Japanese — when scientists in Japan took little interest and he could see no way of synthesizing the predicted molecule.

"That was the worst day of my life," said Osawa in an interview last year with

Chemical Intelligencer, recalling the appearance of the *Nature* paper. "This was my baby, and now someone had stolen it."

Osawa, now at Toyohashi University of Technology, applauds the Nobel prize-winners' recognition in that paper of the wider consequences of their discovery.

He has been given credit by Harry Kroto, who in a 1991 paper, describes Osawa's 1970 paper as "a most imaginative and prescient paper". Kroto and his colleagues were unaware of Osawa's work at the time of their 1985 discovery because it was only published in Japanese. Osawa "greatly regrets" not publishing his 1970 work in English.

David Swinbanks

New post saves science and society course at Stanford

Washington. The Science, Technology and Society (STS) programme at Stanford University, under threat earlier this year when the School of Engineering was unable to find a tenured faculty member to run it (see *Nature* 381, 183; 1996), appears likely to be saved following a decision by the university to create a new tenured position.

The move will ensure the continued availability of degrees majoring in STS at Stanford, although responsibility for the popular undergraduate programme will move from the School of Engineering to the School of Humanities and Sciences.

Three social science departments — economics, political science and sociology — have each been asked to put forward a ranked shortlist of candidates for the new position. The successful applicant will have tenure in one of these departments, and will also chair the interdisciplinary STS programme.

University officials say that much of the programme's recent problems stemmed from the difficulty of persuading senior department staff to spend up to half their time running an interdisciplinary programme within an incentive structure that chiefly rewards activities inside their own departments.

The university senate agreed to make the extra tenured position available after protests from students, who argued that the programme was important in building links between engineering, science and social

science at the university.

Stephen Haber, associate dean for social science in the School of Humanities and Sciences, says: "There is wide recognition that STS is an important programme. The issue has been who should direct it." Haber adds that the three departments will compile shortlists of applicants by 1 November, and that he hopes to appoint "a senior scholar from within the discipline" of STS before the end of this academic year.

But Robert McGinn, a non-tenured professor who currently runs the programme in the School of Engineering, says that the university will not enrol new students majoring in STS until the appointment is filled. He says that the departments will face "a tricky situation" in filling the slot, since they will be accepting someone for a rare tenured position in their own departments whose main interest will lie outside, in the interdisciplinary programme.

McGinn also expresses concern that the programme will lose its traditional emphasis on the impact of technology, and therefore its relevance to engineering students.

But Haber promises that the programme "can maintain its emphasis on technology and society". He says that a significant proportion of the teaching is already taking place in the School of Humanities and Sciences, and he hopes that people from the engineering school will continue to play an important role when the directorship is moved.

Colin Macilwain

Demonstrators call on Yeltsin to honour his budget pledge

Moscow. Labour unions representing Russian scientists staged demonstrations in several cities last week demanding full and immediate payment of government funds promised in the 1996 science budget, 30 per cent of which remains unpaid.

Around 2,000 scientists gathered in Triumph Square in Moscow, urging the government to honour its pledge, signed by President Boris Yeltsin, to reserve 4 per cent of government spending for science. Demonstrations were also held in the cities of St Petersburg, Novosibirsk, Ekaterinburg, Nizhniy Novgorod and other places that host Russia's science establishments.

The demonstrators were joined by Vladimir Strakhov, 64, director of the Earth Physics Institute of the Russian Academy of Sciences, of which he is a senior member. Strakhov has been on hunger strike since the end of September with Igor Naumenko-Bondarenko, chairman of the institute's trade union committee. Both claim that the government's cavalier treatment of Russian science is causing irreparable damage.

Strakhov says 30 of the Earth physics institute's 900 employees have left the country. "The departure of only ten such members will cause a loss on a national scale," he adds. What he describes as "mafia capitalism" is "crunching the bones of Russian science". Strakhov says that science "has reached its death throes" but the government either has not noticed or has decided to ignore the situation.

But Strakhov is a lone voice in the upper echelons of the Russian Academy of Sciences. A recent meeting of the academy's presidium, also attended by directors of the academy's institutes, concluded that, under the country's present economic circumstances, it was "sinful" for academy members to complain, particularly as other state employees, such as teachers and other 'cultural workers' had received only 40 per cent of funds allocated in their budgets.

The academy praised the government for cancelling debts incurred by its institutes, giving these priority over social expenditure. It also welcomed the government's decision to find an additional 80 billion roubles (US\$14.7 million) to cover its debts to scientists, and to promise to increase this sum by two and a half times by the end of the year.

The presidium declined to make additional demands on the government, on the grounds that the state had done what it could for science. But Strakhov and his supporters remain angry. Strakhov complains about "the criminal pittance awarded to the overwhelming majority of scholars. He says he is determined to launch a Russian-wide strike next month.

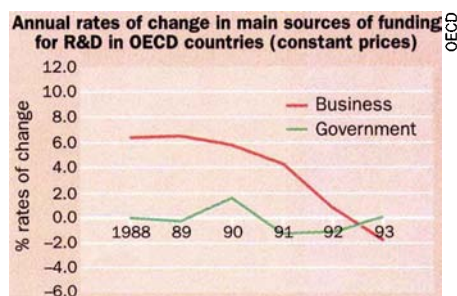
Carl Levitin

State research spending stops declining

London. Governments in the developed world are cautiously increasing their spending on research and development (R&D) now that the effects of recession are beginning to wear off, according to a new report from the Organization for Economic Co-operation and Development (OECD).

But the report, *Science, Technology and Industry Outlook 1996**, also says that science funding is increasingly assessed according to economic and commercial criteria. This creates pressure for more applied science projects. There is "a growing unwillingness to invest in science purely 'for the sake of science'", says the report, suggesting that governments may want to intervene to ensure that "the right balance of R&D is being performed".

The report is the first in a series of two-yearly studies published by the 27-member OECD. It shows that total spending on R&D as a share of gross domestic product in 1993 was 0.2 per cent short of the 2.4 per cent achieved by member countries in 1990.



While governments are spending more on research, funding by industry continues on its downward slope. But the report points out that the business sector still accounts for 60 per cent of R&D spending.

The report shows that a record amount of industrial R&D is now linked to the service industries, such as computers and telecommunications, at the expense of manufacturing-related research. Service-sector research leapt from 1.3 to 18.2 per cent of business R&D funding in the United Kingdom between 1981 and 1993. □

Genetic resistance spreads to consumers

Washington. A small Iowa company that produces a test able to detect genetic alterations in crops is being deluged with calls as wary European consumers and retailers react to the news that genetically modified corn (maize) and soybeans are being harvested in the United States for the first time.

"There has been tremendous interest," says Jeff Wells, who helped to found Genetic ID only two months ago. "It's the European export trade that's interested," he adds, pointing out that it's "not just health food people [but] general consumers".

According to Wells, orders from corn and soy brokers on both sides of the Atlantic have already pushed his firm to capacity. The firm sells tests that can detect the genetic alterations in engineered soybeans and corn produced in the United States by Monsanto and Ciba-Geigy respectively.

Both crops are genetically altered to resist herbicides made respectively by Monsanto and AgrEvo, a Frankfurt chemical company. The corn is also altered to resist the European corn borer, by insertion of a gene from a bacterium called *Bacillus thuringiensis*, which makes a protein toxic to the pest.

Both the soybeans and the corn have been passed as safe by the governments of the United States, Canada and Japan. The European Union (EU) has approved the soybeans for import and processing, but similar approval of the corn is bogged down in committees in Brussels.

Genetic ID's new tests allow middlemen who funnel exports of corn and soy from the United States to European clients to guarantee that the crops are free of genetically altered product — a guarantee that many European retailers are now demanding.

In a separate move, one of the largest US soybean processing companies has begun segregating soybeans. Last month, Central Soya Co of Fort Wayne, Indiana, acknowledged that it had barred deliveries of genetically engineered soybeans to one of its seven US granaries.

According to surveys, up to 85 per cent of European consumers would shun genetically altered foods if given a choice. EuroCommerce, a trade group representing one-third of EU food wholesalers and retailers, demanded here last week that US exports of Monsanto's soybeans be labelled.

At the same time, an international coalition led by consumer activist Jeremy Rifkin announced a boycott of the corn and soybeans, focusing on ten companies which the coalition will pressure to guarantee that certain foods are free of the genetically-altered crops and their derivatives.

The campaign comes amid rumblings of a possible 'corn war' between the United States and Europe. Even as European consumer resistance has mounted to the EU's April decision to allow the import of the soybeans, European countries have divided over whether to allow the import of Ciba-Geigy's genetically engineered corn.

After the EU's Council of Environmental Ministers failed to approve the corn in June, the issue was referred to three scientific committees of the European Commission (EC). It is not known when these committees — on food, animal nutrition, and pesticides — will make their recommendations.

But the first corn grown from the genetically engineered seeds is already being harvested in the United States, and exports to Europe usually start in November. Unless the EC issues rapid approval, imports would violate EU rules.

Ciba says that the corn — 1 to 2 per cent of the expected US harvest — cannot practically be separated from non-genetically engineered corn. This raises the possibility that Europe would need to ban all US corn imports — valued at US\$500 million in 1995 — to bar the genetically-engineered variety.

Without a European decision soon, "there could be serious trade repercussions", says a senior official at the US Department of Agriculture (USDA).

US officials have dismissed as irrational European concern about the safety of both the corn and the soybeans, and have called impracticable demands for their segregation and labelling.

Paul Drazek, senior trade adviser to agriculture secretary Dan Glickman, recently called the soybeans "absolutely safe", and said segregation or labelling make "no scientific sense". At meetings last week in Brussels, Drazek pressed EU officials to speed approval of the Ciba corn, afterwards stating that "there are a lot of people in Europe who understand the importance of these new technologies to world agriculture".

Arnold Foudin, deputy director of biotechnology, biologics and environmental protection at USDA, said last month that the Ciba corn presents "an opportunity to displace a system [of pesticide use] which may be 100 times more damaging".

But European countries have raised concerns about the corn. Last spring, Austria, Denmark and Sweden raised worries about its environmental effects, including the development of pest resistance to the toxin produced by the corn.

They also worried that the corn's herbicide resistance gene might jump to neighbouring weedy relatives. And they complained that the imports ought at least to be labelled as genetically engineered. Consumer activists argue that labelling would protect against unidentified allergens possibly carried by both crops.

Britain raised separate concerns over the fact that the corn carries a marker gene conferring resistance to beta-lactam antibiotics including ampicillin, which is used widely in people and animals. The UK's Advisory Committee on Novel Foods and Processes (ACNFP) has declared that this poses an "unacceptable risk", as bacteria in the guts of animals eating the unprocessed corn could take up the gene.

Ciba disputes all these charges. "We know, and the most stringent review processes from [three] nations confirm, that Bt corn is safe for humans, animals and the environment," says Rich Lotstein, director of regulatory affairs at Ciba Seeds, a division of Ciba-Geigy.

Some prominent scientists, whose opinions were solicited by the company, concur. "There is no scientific data" indicating that DNA could jump from a food to a microbe in the gut of an animal, wrote Roger Beachy, the head of the Division of Plant Biology at Scripps Research Institute in La Jolla, California. He concluded that transgenic foods "pose no risk to the public, nor to the farm animals for which they serve as food".

A Monsanto spokeswoman said that the firm's soybeans have been found safe "around the world", including Japan, Mexico and Argentina. "These are just soybeans, the same as other soybeans, and that's why no special handling or labelling is required," said Karen Marshall.

The huge, St. Louis-based company is evidently undeterred by the controversy. Even as Greenpeace activists sprayed one of its Iowa soybean fields with a giant pink 'X' last Thursday, Monsanto announced that it is considering selling off its \$3.7-billion chemical unit, leaving it better placed to concentrate on its lucrative agricultural biotechnology products — including soybeans. The company's stock rose \$0.50 on the news, to \$41.13. It has gained almost 70 per cent this year.

Meredith Wadman

IMAGE
UNAMIGABLE
FOR A COYALGET
FOR REASONS
REASONS

Growing problem? Some European consumers are worried about the safety of genetically altered corn.

Adam Hart-Davis/SPL

Funding row delays Brussels BSE research

Paris. A fierce row has broken out between officials of the European Commission in Brussels and the French government over the financing of a proposed increase of ECU50-million (US\$62.5-million) in the European Union's (EU) research effort into bovine spongiform encephalopathy (BSE) and Creutzfeldt-Jakob disease (CJD).

Research ministers of the EU member states last week approved the proposal for the new ECU50-million programme. But

they failed to agree on a request from Edith Cresson, the EU's research commissioner, that the money should be taken not from existing EU research funds, but from the ECU200 million remaining in the commission's internal budget. As a result, a decision on how the programme will be financed will not be taken until the next meeting of research ministers on 5 December.

Cresson argued that extra money was needed because there was not enough

money left in the commission's 1997 life sciences budget to cover the ECU50-million programme, according to her spokesman, Jean-Christophe Filori. He says that this proposal was opposed only by "France, France and France", in the person of François d'Aubert, the French secretary of state for research.

The result, claims Filori, will be an unnecessary delay in the launch of the programme. "It's absolutely ludicrous that at a time when ministers are writing cheques of billions of Ecus to support the beef market [following the BSE crisis], they are being petty about ECU50 million for research."

But his claims are strongly denied by the French ministry of research, which accuses the commission of "disinformation". A spokesperson for d'Aubert says that the proposal agreed last week for an increase in the EU's research effort in BSE/CJD was itself originally put forward by d'Aubert in a letter to Cresson in June, and that d'Aubert later specifically asked the EU's Irish presidency to have it put on the agenda of the meeting of research ministers. "We are very pleased that the meeting approved the French proposal," he says.

The spokesperson also contests Cresson's claim that the research budget does not have enough funds left to cover the costs of the proposed programme. He says that d'Aubert in fact favoured taking the money from the research budget, because it would speed up the process. The transfer of money between EU budgets involves the "most complicated and drawn out budgetary procedures the commission has", he says, arguing that in contrast, diverting money from the research budget to the proposed BSE/CJD programme "could be done in a matter of days", as it would only require a vote by a programme committee.

Moreover, d'Aubert's spokesperson claims that all the member states — and not just France — opposed Cresson's proposal for funding of the programme. He alleges that the research directorate is stirring up trouble to score political points in its bid to obtain an injection of extra money for the agreed ECU13.1-billion EU research budget. Indeed, Cresson has hoped to obtain an increase of ECU700 million in the budget. But, ironically, the commission's overall budget has become strained by the BSE crisis, and a meeting of EU finance ministers this week was expected to reject her request.

A spokeswoman for the UK Department of Trade and Industry agrees that France was not alone in refusing to agree to Cresson's funding proposal last week. Britain's view was that the commission "needed to verify how much it had in its research coffers" before any decision could be made at the next meeting of research ministers in December, she says.

Declan Butler

Panel urges precautionary approach

Paris. The UK government's optimism that bovine spongiform encephalopathy (BSE) will die out around the end of the decade (see *Nature* **383**, 209; 1996) received a further blow last week. A panel of scientists set up by the European Commission to recommend priorities in research into transmissible spongiform encephalopathies warned that the risk that the disease would become endemic cannot be discounted, and that research is urgently needed to anticipate this possibility.

The panel, chaired by Charles Weissmann, from the Institute of Molecular Biology in Zurich, pointed out in its report to the commission that predictions that the disease will be eradicated are based on assumptions that might not "hold true", while the agent that causes the disease might mutate, or be transmitted in ways that are not yet known. As a result, it proposes a detailed programme of fundamental research into the transmissible spongiform encephalopathies.

Weissmann added that action should be taken as if transmission of the disease between cattle and humans "were true". The report calls for a review of the epidemiology of BSE, and of the infectivity of cattle tissues and the products in which they are used, arguing that the mouse assays used previously to evaluate infectivity "are very insensitive."

Continuing in the same precautionary vein, the report calls for the determination of "the contamination level of meat with brain/spinal cord after standard butchering procedures." Similarly, it calls for a review of maternal risk factors in the transmission of BSE in cattle, and the determination of the infectivity of "placenta, blood, colostrum, milk, and excrements of BSE cattle".

The report, which was discussed at a press conference in Brussels last week by Weissmann and Franz Fischler, the European Union's agriculture commissioner, also calls for a series of studies

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Weissmann (right) and Fischler outline recommendations for research into BSE.

to assess the risk of BSE to humans.

Surveillance of BSE and CJD should be extended throughout Europe, according to the report, adding that CJD epidemiology should "monitor for linkage to consumption of cattle and sheep products". The panel's reference to sheep products will be enough to send shudders through the commission's agriculture directorate.

Indeed, Weissmann's report makes explicit the fact that sheep may also pose a risk to human health. "It is possible that BSE originated as a (rare) spontaneous event in cattle and that the BSE agent is different from the scrapie agent. If so, then BSE may have been transmitted to sheep via bone and meat meal from cattle and BSE infected sheep could transmit the agent to man." As a result, it recommends EU-wide surveillance of scrapie, and studies to identify the infectious agent in recent UK cases of "sheep 'scrapie'".

Again assuming the worst-case scenario, the report calls for an accurate assessment of any past risk to which humans were exposed, including an analysis of the use of offal in various foods, and the consumption of these by various "age groups and socio-economic criteria." A diagnostic test for CJD/BSE is a priority, it adds.

D. B.

AP/Thierry Charlier

Germany stunned by institute closures

Munich. Germany's Max Planck Society (MPS) shocked the country's scientists by announcing last week that it plans to close four of its institutes in the western part of the country to meet government demands for staff cuts. Further closures are likely.

Hubert Markl, president of MPS, says the closures are needed because the federal government, which provides half the research organization's funding, unexpectedly ordered a 5 per cent cut in staff, equivalent to 320 posts, earlier this month. This comes on top of the 420 staff that the MPS had already been told to shed by 1998.

Following Germany's reunification, the society has had to streamline western institutes to allow new ones to be established in the east, and it has already made cuts to achieve this. As a result, the new staff savings cannot be spread across all the institutes without affecting efficiency, according to Edmund Marsch, MPS deputy general secretary. He warns that the closure of the four institutes will only result in half of the required staff savings, and further closures will therefore follow.

Three of the institutes earmarked for closure are the Institute for Aeronomy in Lindau, Niedersachsen, with 185 staff, the Institute for Biology in Tübingen, Baden-

Württemberg, with 93 staff, and the small Institute for Human Ethology in Andechs, Bavaria. The fourth has not yet been named because its director is currently abroad.

Marsch stresses that the closure decisions were not based on the scientific performance of the institutes. A significant factor in deciding which institutes to target, he says, was the impending retirement of their respective directors. The MPS often chooses to close a department after the retirement of its director. But this is the first large-scale closure of whole institutes.

Closure of the Institute for Human Ethology will be almost immediate. But the last of the four directors of the Lindau institute will retire in ten years' time, and in Tübingen the last director is due to retire in 2004, so these institutes will be wound down slowly. That will be hard on scientists who will feel they are working under a death sentence, says Peter Overath, scientific director at the Institute for Biology.

The institutes were unaware that closures were imminent. Only two weeks before the announcement, says Overath, the Institute for Biology had hosted a symposium at which candidates to replace two of the retiring directors delivered lectures as part of the normal selection procedure.

A spokesman for the Institute for Aeronomy says the institute had hopes of being a major player in Rosetta, the cornerstone mission of the European Space Agency (ESA) which will fly into a comet nucleus, and this could now be thrown into doubt.

According to David Southwood, former chairman of ESA's Science Programme Committee, the loss of the institute's expertise in high-quality instrumentation would be serious not only for Germany but for the many other countries that collaborate with it on space science. But Marsch says that the work at the institute is no longer a priority for the society.

Markl's closure plans will be discussed on Friday 18 October by the MPS's technical committees. The society's senate will vote on the overall plan on 22 November, and the individual closures will be put to a vote of a senate meeting next spring. Meanwhile targeted institutes are preparing their defence, in the hope that the society may be persuaded to change its mind. But Markl has said from the beginning of his term as president that he will not shy away from difficult decisions about institute closures in west Germany (see *Nature* 381, 357; 1996).

Alison Abbott & Quirin Schiermeier

Clinton hooks on to science as 'a bridge to the future'

Washington. Science has emerged as a prominent theme of President Bill Clinton's campaign speeches as he travels around the country promising to build "a bridge to the 21st century" if — as widely expected — he is re-elected to a second four-year term on 5 November.

Last week, for example, Clinton praised the Human Genome Project, cancer and AIDS research, space science and the Internet, in remarks at campaign stops around the country. He and Vice-President Al Gore accused Bob Dole, the Republican candidate, of planning across-the-board cuts in research funding and argued that Dole's plan to abolish the Department of Energy would endanger several national laboratories.

Speaking in Knoxville, Tennessee, Clinton pointed out that all five US winners of this year's Nobel prizes in physics and chemistry had been funded by the federal government through the National Science Foundation (NSF). "Cutting back on research at the dawn of a new century would be like cutting our defence budget at the height of the Cold War," he said. "We must not do it and we will not do it."

Clinton singled out the Human Genome Project as "one of my favourites", promising his audience that "it won't be too many years before parents will be able to go home

from the hospital with their new-born babies with a genetic map in their hands that will tell them: 'here's what your child's future will likely be like'.

Clinton has rarely mentioned science during his first four-year term, and political commentators suggest that his apparent interest now is merely part of a strategy to make 73-year-old Dole look old-fashioned. The Dole campaign has not responded directly to the attacks.

Gore, worried that the loss of his home state of Tennessee could disrupt his rise as the likely Democrat presidential nominee in 2000, told the Knoxville audience that Dole had "proposed measures that would clearly shut down" the nearby Oak Ridge National Laboratory. "President Clinton and I will not let that happen," he pledged.

But the prospects for serious discussion of the science issues that divide the two camps — such as the future of the Department of Energy — receded with the cancellation of plans by the American Association for the Advancement of Science (AAAS) for a debate on them in Washington.

Mary Good, assistant secretary at the Department of Commerce, and Robert Walker, the retiring chairman of the science committee in the House of Representatives, had been lined up to represent the Clinton and Dole campaigns respectively in the

forum on 22 October. But the AAAS cancelled it late last week.

With Clinton's 15-point lead in opinion polls holding steady since early summer, attention is beginning to drift from the campaign to the nature of a second Clinton administration. Jack Gibbons, the president's science adviser, is expected to retire, and his choice of Good to represent Clinton in the aborted debate reinforced her status as front runner to succeed him.

A former chemistry professor and industrial research manager, Good's energy and directness would, in the eyes of science lobbyists, more than compensate for her lack of top-rate scientific credentials. Other names in the ring are Ernest Moniz, a physicist currently assisting Gibbons, James Baker, the administrator of the National Oceanic and Atmospheric Agency and — curiously — Dan Goldin, the administrator of NASA.

Neal Lane has a six-year term as director of the National Science Foundation, ending in 1998, which he would like to complete, and Harold Varmus, director of the National Institutes of Health, is also expected to stay on. Hazel O'Leary, the energy secretary, will almost certainly leave and, despite efforts on Capitol Hill to rally support for Charles Curtis, her well-respected deputy, to succeed her, a political appointment from outside the department is likely. **Colin Macilwain**

Funding cuts put pressure on peer review

Capri, Italy. Cuts in research funding, and the resulting increase in competition for funds between scientists, have exposed the shortcomings of the peer review system. But it remains the best system in operation, according to an international 'consensus' conference on research assessment held in Italy last week.

The conference, on the island of Capri, near Naples, was held under the auspices of an informal group of representatives from G7 countries. The scientists and policy-makers from three continents who attended agreed that the peer review system needs to be improved to fit new circumstances.

They also agreed that objective criteria, such as citation data, can be helpful in assessing research departments, but should only be used very carefully when assessing individuals. This was particularly true for the academic community of the host country, Italy, which is struggling to institute wide-ranging reforms, but has little confidence in its current procedures for reviewing grant applications and for career promotion.

Much of the current pressure on peer review results from oversubscription to research programmes, with the result that the success rates of applications have fallen to levels widely considered unhealthy. "This not only strains the time of reviewers, it also strains the concept that the system is operating fairly," said Susan Cozzens, a policy officer at the US National Science Foundation.

Cozzens said low success rates tend to favour solid, but low-risk, applications. Others argued that this situation also undermines the value of peer review, as the difficulty of choosing objectively between closely-ranked top-quality applications can make the selection of successful projects little more than a lottery.

Several participants pointed out that grant agencies are acutely aware of the strain on reviewers. Many leading scientists are unhappy with the drudgery of wading through stacks of low-quality grant applications, and frustrated that applications they rate highly remain unfunded because of lack of money.

Some funding agencies are experimenting with pre-screening applications to weed out 'no-hopers' at an early stage. But pre-screening, with its spectre of 'assessment by administrators', is unpopular in much of the scientific community, and can add to a general perception of unfairness.

The conference agreed that confidence in the fairness of a peer-review system can be increased by the appropriate use of bibliometric data, such as publication rates and 'impact factors', which measure citation rate. But participants acknowledged that such data should be used carefully, particularly in assessing individuals, and should never replace human judgement entirely.

It is a sensitive issue. Many researchers feel that the mathematical detachment of citation data implies a certainty that is attractive, but is not always warranted. The conference agreed that such data can be misleading, for example on multi-author papers, as a laboratory chief or supplier of biological material may routinely add his or her name to papers to which they have made no scientific contribution.

Indeed the conference was told by Eugene Garfield that citation indices had never been intended for evaluation. Garfield, who developed the concept of impact factors as a measure of a journal's influence, is founder, and now emeritus chairman, of the Institute for Scientific Information in Philadelphia.

Garfield said that he had been widely

depicted as "a Frankenstein" whose monster — the impact factor of an individual paper — had an enormous potential to be misused "in the hands of uninformed users".

Pressure to rely heavily on publication-based indicators is particularly strong in Italy, where many grant-giving agencies and charities distribute small amounts of money evenly, to avoid upsetting unsuccessful



applicants, and the system of academic promotion through national competitions, or *concorsi*, is widely criticized.

Many Italian scientists feel that the lack of formal selection criteria has allowed personal contacts to play too important a role in the allocation of professorships and assistant professorships. Those at the conference described how the academic community is split over how much weight should be given to citation data in assessing an individual's research performance.

Some feel that such data should at least be used to eliminate individuals with low citation ratings at an early stage. Complex formulae for weighting the value of each publication have been suggested as ways of compensating for the weaknesses of citation data — for example, scaling down the impact factor of a multi-author paper or a review, which normally have higher impact factors than a research paper.

Proponents argue that such formulae would allow citation data to be used fairly. But others fear that excessive reliance on such data could allow the merits of a first-class scientist, who has published little, to be overlooked.

The conference agreed in one of a series of formal statements that citation indices should only be used as one of a number of different performance indicators. It said that an assessment committee should be able to give a high rating to a 'low impact' individual — and vice versa — provided that it is able to give an open justification of its decision.

The conference also agreed that citation indices are too crude to distinguish between closely-ranked grant applications. As a result, complex formulae intended to quantify the merit of individual publications are unnecessary.

Alison Abbott

Europe confronts barriers to mobility

Munich. The research ministers of member states of the European Union (EU) have welcomed a Green Paper published by the European Commission that identifies the legal and cultural difficulties faced by students and young researchers who wish to train in another EU country, and suggests how they might be overcome.

The document highlights the problems that the commission has experienced in running its many research and training programmes intended to promote mobility within the scientific community. Many of the difficulties are rooted in the differences in taxation and social security laws between member states.

Until recently, for example, the take-home value of the commission's

research fellowships, after the deduction of income and social security taxes, varied so much that researchers in some countries found themselves virtually on the breadline. Such problems, on top of language and cultural differences, discourage mobility of scientists.

The Green Paper suggests that a legal framework should be established for students and research trainees to address such problems. This would include the recognition by all member states of a European research grant which would be exempt from tax. The commission also proposes that mobility training programmes should include language lessons for those moving to another country.

A. A.

Mount Graham telescope brings further partners into sight

Washington. Ohio State University is planning to rejoin the Large Binocular Telescope (LBT) project on Mount Graham in Arizona, which it left in 1991 after protests against the environmental impact of the project.

The university is planning a \$6.4-million contribution to LBT, buying it a one-eighth share in the project. The much-delayed facility is now under construction near the summit of Mount Graham, and a statement from Ohio State said that the environmental and land-use issues surrounding the project "have been satisfactorily resolved".

Existing partners at the LBT — the University of Arizona, the Research Corporation, and an Italian consortium — have enough money to equip the telescope with only one mirror. But if negotiations with Ohio State University and a German consortium are successfully completed, the two-mirror project will be fully financed. It could be in use in 2001, the statement said. □

Denver can upgrade observatory

Washington. After a year-long environmental assessment, the United States Forest Service has issued a decision notice and special use permit allowing the University of Denver to upgrade its facilities at Mount Evans Observatory near Georgetown, Colorado. This permission includes provision to add a 4-metre-class telescope to the observatory's existing facilities, about 4,000 metres above sea level. The university has issued a call for prospective academic and financial partners for the project. ■

Apology for accused dean

London. The rector of a Canadian university that wrongfully forced out one of its senior staff following allegations of financial misconduct has publicly expressed regret at the pain and loss of reputation suffered by the researcher, M. N. S. Swamy, and his family. Frederick Lowy, rector of Concordia University, Montreal, has acknowledged the university's mistake and confirmed that "there is no evidence of mismanagement or misuse of any other funds under Dr Swamy's control".

In 1994, Swamy, then dean of engineering and computer science, and two other researchers had their research grants frozen following allegations of financial irregularities. Swamy was asked to retire early along with one other colleague. The third researcher was asked to resign. The allegations were later found to be without foundation. But the university, despite clearing the three of any misconduct, was unable to offer them faculty positions (see *Nature* **381**, 104; 1996). Swamy, however, has now been nominated for the position of professor emeritus. ■

French chemical charges

Paris. A French historian last week alleged that a French chemical company, Ugine, may have produced Zyklon B, the gas that was used in the Nazi gas chambers of the second world war. The claims were made by Annie Lacroix-Riz, professor of modern history at the University of Toulouse, and a member of the French communist party. In support of her claims, Lacroix-Riz cites the creation in 1941 of a Franco-German company — Duferrit-Sofumi — by Ugine and Degesh, a company which was itself partly owned by I. G. Farben, the chemical group that produced Zyklon B. ■

Biomedical institute expands

Boston. A new wing has been dedicated at the Massachusetts Institute of Technology's Whitehead Institute for Biomedical Research. The \$35-million expansion, which was completed after a four-year construction effort, adds 20 new laboratories, increasing by

more than 45 per cent the amount of space available for research and training at the institute, which is playing a key role in the Human Genome Project.

The new wing will enable the institute to expand its biologic containment facilities for infectious disease research, double the size of its facilities for studying mouse models of human disease, and create a new Center for Structural Biology that will promote the field of molecular medicine. The Whitehead Fellows Program will also grow as a result of the research space opened up by the new wing. "The new facilities will ensure that our young researchers continue to have the tools and resources they need to pursue novel ideas," says Gerald Fink, director of the Whitehead Institute. □

Award for promoting science

London. Sweden's Right Livelihood Foundation has honoured an Indian organization that has adopted a pioneering approach to public understanding of science in the southern state of Kerala. The three recipients of the 1996 Right Livelihood Award — sometimes described as the 'alternative Nobel prize' — include the Kerala Sastra Sahithya Parishat (KSSP), the people's science movement of Kerala.

The KSSP communicates science by simultaneously emphasizing its impact on society. KSSP now has 60,000 members, including 10,000 teachers. It is widely considered to have contributed to Kerala's high levels of adult literacy — 91 per cent — and life expectancy — 71 years — compared with the Indian average of 52 per cent and 60 years respectively. □

Biologist for US agency

Washington. The former second-in-command at the short-lived National Biological Service (NBS) has been tipped to become the first chief biologist of the reorganized US Geological Survey (USGS) in Reston, Virginia. Dennis B. Fenn, a soil scientist who recently served as the NBS deputy director, will head the new Biological Resources Division at USGS (see *Nature* **382**, 658; 1996). The survey's fourth and newest division was formed from the remains of the NBS, which Congress abolished as an independent agency last year. Before joining the NBS, Fenn was chief scientist at the National Park Service. □

Astronomer backs 'ET research'

Paris. Sir Martin Rees, Britain's Astronomer Royal, last week said he supported a "modest" programme to search for intelligent extraterrestrial life. Rees was speaking after delivering a lecture on cosmology at the British Embassy in Paris, organized by the British Council as part of France's science festival, *Science en fête*.

Although radioastronomers felt searches for extraterrestrial life "diverted their time", Rees argued that public interest in the area was "enormous". Rees himself declared that he had "an open mind" as to existence of life on other planets. Meanwhile, the Royal Greenwich Observatory this week called for an international effort to study the threat that other sorts of extraterrestrial bodies, such as asteroids and comets, might collide with the Earth. □

Population growth decelerates

London. Global population growth appears to have slowed down, but the steep rise in the numbers of elderly people will continue, according to new forecasts from the International Institute for Applied Systems Analysis (IIASA), a private research organization based in Laxenburg, Austria.

Estimates presented at the United Nations population conference in Cairo two years ago projected a doubling of world population to 12 billion within the next century. But the IIASA analysis, published earlier this month, says that the continuing decline in fertility strongly indicates that this 12-billion figure will not be reached. □

Genetics and ethics in China

SIR — In response to the article on the Chinese Maternal and Infant Health Care Law (*Nature* 383, 204; 1996), I should like to clarify the position of the International Genetics Federation (IGF) on the relation between the so-called 'eugenics law' and the 1998 International Genetics Congress, to be held in Beijing.

One important task of the IGF is to sponsor and oversee the administration of international genetics congresses. At the 1993 Birmingham Congress, the representative council of the IGF, composed of delegates from genetical societies worldwide, voted to hold the next congress in Beijing.

Subsequently, when the new Chinese law was passed, the executive council of the IGF decided that the best way to deal with the genetical and ethical issues that are inherent in it was to encourage a full debate among geneticists. The Beijing Congress seemed (and still seems) the best venue for such a debate. The executive therefore asked the local organizing committee in Beijing to include a symposium on eugenics in the programme. The committee agreed and the symposium will be organized by Dr Robert Haynes, who was president of the highly successful 1989 congress in Toronto. (Haynes has never held a position in the IGF, as you claim.)

You reported on a statement issued by eminent human geneticists at the 9th International Congress of Human Genetics (Rio de Janeiro, 1996) calling for a delay in implementing the Chinese law, and a moratorium on the drafting of similar laws in other countries pending discussion by geneticists. This statement is perfectly in accord with the notion of using the Beijing Congress as a platform for an international debate on these serious issues.

You also reported the result of a ballot circulated to members of the British Genetical Society on a motion proposing suspension of the society's affiliation with the IGF as a protest against its decision to hold the next congress in Beijing. The result was that 56 individuals voted for and 44 against the motion (not 60 to 40 as you reported).

This result must be treated cautiously. First, the sample size was small (about 6% of the membership). Such a small return could be interpreted in several ways, but the danger is that the sample is not representative. Second, before considering disaffiliation from the IGF, the Genetical Society should note that the IGF is active in other aspects of genetics at the international level. It provides small grants for local conferences in poor countries, supports travel by young geneticists to international genetics congresses, and has initiated schemes to promote worldwide

education in genetics and to provide partnerships between geneticists in rich and poor countries. All these would be affected adversely by withdrawal of support of one of the world's largest societies.

Third, the protest seems misplaced. The IGF can act only through the constraints of its constitution and did so. If changes in the constitution are required, then member societies can propose them and vote for them. Furthermore, impeding the success of the Beijing Congress would hurt participating groups that seem unlikely to be the true targets of the protest. Chinese and other geneticists would be deprived of an important opportunity to hear and meet the world's best scientists.

Genetics and ethics have always been intertwined. We should use the international gathering in 1998 to explore and debate this relationship in the complex and controversial area of eugenics, and together arrive at comprehensive set of recommendations to governments worldwide. For such recommendations to carry the full weight of authority, it is important that all genetical societies attend and participate.

A. J. F. Griffiths

(Secretary, International Genetics Federation)
University of British Columbia,
Vancouver, BC, Canada
e-mail: agriff@unixg.ubc.ca

Shells suitable

SIR — Daedalus proposes that chitin fibres and bulk chitin polymer might be produced from crab shells (*Nature* 382, 498; 1996). This has already been done by Professor Seiichi Tokura at the Graduate School of Environmental Earth Sciences of Hokkaido University, Sapporo, Japan. Tokura routinely produces both very strong chitin fibres and thin chitin sheets from crab shells.

The fibres are very useful for stitching after surgical operations, because they are completely absorbed by the human body and do not have to be physically removed. Similarly, the chitin sheets serve as bandages for bleeding wounds. It is often difficult to remove ordinary bandages, which stick to a wound because of clotted blood, but the chitin bandages are simply absorbed into the body, and more layers can be applied until the wound has healed. There are, of course, many other potential uses of crab chitin, and many of these are being developed by Tokura.

Harmon Craig

Scripps Institution of Oceanography,
University of California, San Diego,
La Jolla, California, 92093-0220, USA

Shiny Sheffield

SIR — I was disappointed by the photograph of Sheffield that you used to illustrate your recent News story on British government plans to cut air pollution by 2005 (*Nature* 382, 743; 1996).

The photograph was clearly taken some years ago. Sheffield is now rightly regarded



In the clear: Sheffield looks like this now.

as one of Europe's cleanest industrial cities and is pioneering new ways of ensuring management and sustainability of air quality. To that end, we have been acting as a pilot for the government's draft strategy that you reported, and we have just received European funding to assist our work.

Frank R. Price

Environmental Protection Service,
Sheffield City Council,
1 Barkers Pool,
Sheffield S1 1EN, UK

Multiple author

SIR — I have been wondering about the law of priority for names of authors of papers. I am one of those relatively rare males who have a problem that is all too common among females — I have changed my name since I started publishing.

One solution is to continue to use the 'maiden' name (J. T. Clark), but having decided to abandon my first surname this is not one for me to adopt. I have tried as far as possible to use both surnames in published articles (J. T. Clark Sellick) to get some form of continuity, but this gets me extracted in indexes as Clark, Sellick and even Clarksellick. Some journals, however, now insist on just first name and surname (John Sellick), so that neither previous surname nor familiar initials provide any continuity.

Can anybody recommend a solution? I was inspired to pen this note by seeing a footnote to a citation that read "Clark is a senior name to Sellick".

J. T. C. Sellick

Montemassi, 31 Regent Street,
Kettering,
Northants NN16 8QG, UK

strings are thought to form, we get the helium analogy — vortices. Bingo.

Meanwhile, your domestic equilibrium has been upset by contemplating our temerity in extrapolating a few millilitres of superfluid to the scale of the Universe. It seems only our duty to put minds at rest and restore conjugal bliss. Right, point out to him — very gently — that when the Universe underwent this transition, it was probably no bigger than a golfball.

C. Bäuerle

Yu. M. Bunkov

S. N. Fisher

H. Godfrin

CNRS Centre de Recherches sur les

Très Basses Températures,

BP 166,

38042 Grenoble Cedex 9, France

G. R. Pickett

School of Physics and Chemistry,

Lancaster University,

Lancaster LA1 4YB, UK

Costing the Earth

SIR — Caldeira¹ criticizes the Economic-Damage Index² for using discounting to compare the monetary value of damages incurred at different times. Although discounting can appear easy to ridicule, one must account for the opportunity cost of resources³.

When investment opportunities exist, society can reallocate social costs over time. Consider Caldeira's second example: if, as implicitly assumed in the example, the value of an extinction V (the maximum value of resources society would willingly sacrifice to prevent it) is independent of the date, number and identity of species lost, and additional investments yield 3% per year social return, the time path of social costs associated with losing 10 species in 78 years can be exchanged for a path where a cost of $10V(1.03)^{-78} = 0.997V$ is incurred today for investment, with the proceeds used to offset the value of the species loss in 78 years. In these conditions, losing 10 species in 78 years is not worse than losing one today, as the timing and magnitude of social costs of the former can be made equal to those of the latter if desired.

Conversely, if one refuses to save one species today because the resources so used could be invested to save 10 species 78 years later, should not one subsequently refuse to save the 10 species because the resources could then be reinvested to allow 100 species to be saved in another 78 years?⁴

Discounting is properly applied to the monetary value of changes⁵ and is not applicable when tradeoffs cannot be made between environmental damages and other resources. One might choose to put biodiversity outside the bounds of economic analysis, setting constraints on species loss

that may not be violated regardless of the cost in other resources, but such an approach leaves open the question of how to set the constraints.

James K. Hammitt

Center for Risk Analysis &

Department of Health Policy
and Management,

Harvard School of Public Health,

718 Huntington Avenue,

Boston, Massachusetts 02115, USA

Atul K. Jain

Donald J. Wuebbles

Department of Atmospheric Sciences,

University of Illinois,

105 South Gregory Avenue,

Urbana, Illinois 61801, USA

John L. Adams

RAND Corporation,

1700 Main Street,

Santa Monica, California 90407, USA

1. Caldeira, K. *Nature* **383**, 214 (1996)
2. Hammitt, J. K., Jain, A. K., Adams, J. L. & Wuebbles, D. J. *Nature* **381**, 301–303 (1996).
3. Arrow, K. J. et al. in *Climate Change 1995: Economic and Social Dimensions of Climate Change* (eds. Bruce, J. P., Lee, H. & Haites, E. F.) 125–144 (Cambridge Univ. Press, 1996).
4. Keeler, E. B. & Cretin, S. *Management Science* **29**, 300–306 (1983).
5. Hammitt, J. K. *J. Health Econ.* **12**, 117–120 (1993).

corres@nature.com

Letters submitted for Correspondence can be e-mailed to **corres@nature.com**, but please send only items intended for the Correspondence section to this address.

Does your automated DNA sequencer leave you guessing? If so, chances are it's primarily designed for high throughput sequencing. Why be uncertain of your sequencer's accuracy, when ALFexpress™ is providing researchers with the full genetic stories of their DNA.

ALFexpress: for more accurate readings

In the largest clinical study using automated DNA sequencing, the technology behind ALFexpress proved exceptional (see caption). That's one of many examples of ALFexpress offering unrivaled accuracy during automated confirmatory sequencing. Further, its readings are so accurate that ALFexpress can unambiguously identify heterozygous point mutations—as proven in many clinical research applications, such as analysis of tumor genes and high-resolution HLA typing. What's more, Pharmacia Biotech has dedicated software programs to support these applications.

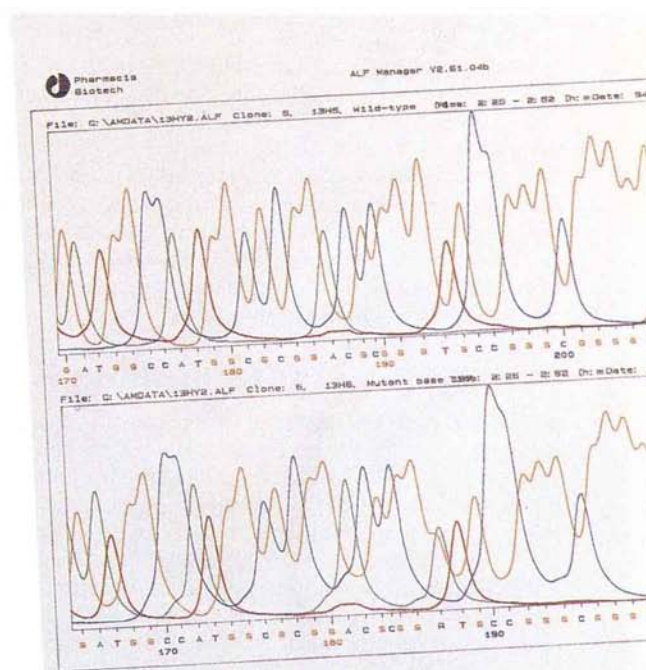
For the full story, call us: 1 (800) 526-3593 from the U.S.; +81 (0)3 3492 6949 from Japan; or + 46 (0)18 16 50 00 from Europe and the rest of the world. Or visit us on the Internet: <http://www.biotech.pharmacia.se>.



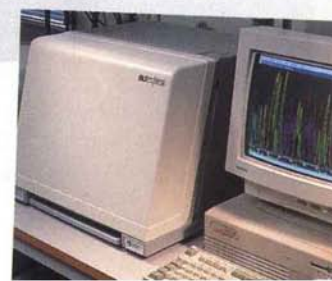
**Pharmacia
Biotech**

Uppsala, Sweden. (And the rest of the world)

READER ENQUIRY NO. 100



The p53 gene from 316 breast cancer patients was sequenced using ALF automated sequencing technology. (Bergh J., Norberg, T., Sjögren, S., Lindgren A., Holmberg, L. "Complete Sequencing of the p53 Gene..." *Nature Medicine* 1995; 10:1029-1034.)



India's declining ranking

SIR — Scientometric studies are useful in understanding the dynamics and performance of science, as well as for policy interventions. Several authors have attempted such quantitative analyses of Indian science¹, but none has analysed the long-term trends in Indian scientific output in relation to world trends.

Our analysis of authored publications in the journals covered by the *Science Citation Index (SCI)* between 1981 and 1995 shows that Indian share in world scientific output has declined by 32%, almost double the 17% decline estimated in *World Science Report*², on the basis of *SCI* data collected for 1982 and 1993. Even at 17%, no other country suffered such a massive decline, except the Commonwealth of Independent States. Consequently, India's world ranking has also declined from eighth position in 1980 to thirteenth in 1995.

How did this happen? A comparison of the annual publication output of India and of the world reveals that Indian publications declined by about 15% between 1981 and 1984 and stagnated thereafter, whereas world output grew by about 22%. It can be argued that this decline is a reflection of the reduced coverage of Indian journals in *SCI*, which fell by two-thirds during the above period³. However, Indian journals accounted for only 35% of the total Indian output even when *SCI* coverage was at its best. Reduced coverage, therefore, accounts for only one-third of the overall decline in the Indian share in world output.

The decline cannot be fully explained by a publishing bias against developing countries either, as other such countries (especially in South-East Asia and China) improved their world share tremendously during the same period (see figure), and there is not enough evidence to suggest any specific anti-India bias. Nor is this decline due to a reduced preference for submission of Indian manuscripts in *SCI* journals (INSDOC survey; Pandale, personal communication). So Indian authors could be facing more rejections than before.

The relative citation rate of Indian research papers, expressed as a ratio of observed citations to expected citations⁴, has also been declining steadily, causing a steep fall in India's rank by citation impact, from 57 in 1985 to 81 in 1989 and probably further since then. This indicates that India is losing out in terms not only of quantity but also of quality. But the practice of judging quality by citations is contentious⁵.

It is therefore clear that the *SCI* data, in spite of their limitations, give a fairly reliable picture of India's declining position in world science. This calls for an in-depth analysis of science management in India. We point out a few possible factors here.

First, on the 'brain drain' of scientists from India to the West⁶, *SCI* rightly identi-



International trends in research publications. The percentage change in the share of different regions 1982-93, based on *World Science Report/SCI* data. (EU, European Union; USA, United States; CAN, Canada; LAM, Latin America; NAF, North Africa; MNE, Middle and Near East; SSA, Sub-Saharan Africa; IND, India; CIS, Commonwealth of Independent States; JAP, Japan; CHI, China; SEA, South-East Asia.)

fies Indian articles only by the address of the first author, and does not take account of Indian authors publishing from abroad. India receives no credit for its scientists' contributions from a foreign country, especially since it failed to retain and support them. Second, as to declining government funding for science in India⁷, although funding has increased in absolute terms, the rate of growth of expenditure on research and development at constant as well as current prices is on the decline⁸.

The 'ageing' of scientific institutions in India⁹ through decreasing recruitment also reduces publishing activity, as older scientists rarely work with their own hands in India. And as young scientists are a floating population of *ad hoc* workers hired for a pittance and fired at whim, they often fail to deliver tangible benefits, especially when there are more lucrative opportunities abroad.

Lack of motivation, a feudal work culture and absence of dynamic and inspiring leadership are equally important. But what is most striking is the total lack of long-term monitoring of Indian science publishing trends as an input to policy and planning.

N. Raghuram

Centre for Science and Environment,
New Delhi 110 062, India

Y. Madhavi*

National Institute of Science,
Technology and Development Studies,
Dr K.S. Krishnan Marg, Pusa,
New Delhi 110 012, India.
e-mail: postmast@csnistad.ren.nic.in

*To whom correspondence should be addressed.

- Gupta, B. M. (ed.) *Handbook of Libraries, Archives and Information Centres in India* Vol. 13 (Segment, New Delhi, 1996).
- Papon, P. & Barre, R. in *World Science Report* (ed. Moore, H.) 8-23 (UNESCO, Paris, 1996).
- Arunachalam, S. J. *Sci. Ind. Res.* **51**, 99-119 (1992).
- Braun, T. et al. *Scientometrics* **33**, 262-293 (1995); **29**, 299-344 (1994); **13**, 181-188 (1988).
- Maddox, J. & Jayaraman, K. S. *Nature* **366**, 611-627 (1993).
- Jain, A., Pruthi, S. Garg, K. C. & Nabi, S. A. *Indicators of Indian Science and Technology* (Segment, New Delhi, 1996).
- Rajeswari, A. R. *J. Sci. Ind. Res.* **54**, 559-570 (1995).

Limited consensus

SIR — John Maddox questions the validity of the 'consensus' sometimes associated with the conclusions of the Intergovernmental Panel on Climate Change (IPCC) (*Nature* **383**, 17; 1996). As he points out, the nature of the scientific process, which is built on debate, argument and disagreement as well as agreement, is not very compatible with the idea of consensus.

The nature of the consensus being pursued by the IPCC, however, is not the achievement of agreement about all of the science. Rather, it is a limited consensus delineating those parts of the science where scientists are in broad agreement and those parts where there remains a great deal of uncertainty and debate. It is this limited consensus — of paramount importance to those formulating policy — that the IPCC Summaries for Policymakers attempt to present as clearly as possible and without any particular political spin.

The very substantial and detailed assessments that form the bulk of the IPCC reports are aimed at providing just what Maddox is asking for, namely a continuing review of the peer-reviewed literature that carries weight in the research community, including an assessment of the uncertainties.

John Houghton

(Co-Chairman, Scientific
Assessment Working Group)
Intergovernmental Panel
on Climate Change,
Meteorological Office,
Hadley Centre, London Road,
Bracknell RG12 2SY, UK

Chain reaction

SIR — There are indications that some scientists have been infected with an agent of unknown nature and origin which leads them to think that there is such a thing as a human food chain¹. The same agent may be responsible for the belief that proteins can be cloned². Without editorial vigilance on your part, the effects of this agent may result in an epidemic of confusion and the further degeneration of our scientific language.

Michael G. Mortimer

Chemical & Biological Sciences,
University of Huddersfield,
Queensgate,
Huddersfield HD1 3DH, UK
e-mail: m.g.mortimer@hud.ac.uk

- Anderson, R. M. et al. *Nature* **382**, 779-788 (1996).
- Oberlin, E. et al. *Nature* **382**, 833-835 (1996).

Correction: Climate change

THERE is no connection between the people listed in Hugh W. Ellsaesser's letter about global warming (*Nature* **383**, 214; 1996) and the companies and institutes shown in parentheses after their names. The institutes should have been listed separately. We apologize for this error. — Editor, *Nature*. □

Odometers for crude-oil migration

Gary H. Isaksen

Trace molecules present in crude oils may hold the key to knowing how far oil has migrated in the Earth's subsurface. If so, improved mapping of migration pathways can aid in the search for new petroleum reserves.

GLOBAL demand for energy continues to grow inexorably. Although use of alternative, renewable sources of energy is increasing, our reliance on petroleum (oil and gas) will continue for many decades. Most giant petroleum accumulations (greater than 500 million barrels of recoverable oil or equivalent gas) have probably already been discovered, as is indicated by their decreasing discovery rates since the early 1970s. Some frontier areas, such as polar regions and remote continental interiors, are not economic with current price forecasts, and exploration and production technologies. Other areas are unattractive because of the instability of the prevailing political regime.

Nonetheless, exploration for giant accumulations proceeds in frontier areas such as Sakhalin Island, Russia, and off the coasts of West Africa and northwest Scotland. Many petroleum companies are also developing technology aimed at finding more petroleum reserves near existing fields. It is in this general context that the paper by Larter *et al.* on page 593 of this issue¹ may prove to be important. The authors provide evidence, both from oil samples and from experiments, that ratios of two compounds present in oils constitute odometers that tell how far the oils migrated in the Earth's subsurface.

The compounds are benzocarbazoles a and c, which are nitrogen-containing isomers. Because of their different structural configuration and the presence of nitrogen, they have different polarities, meaning that they adsorb differently on the mineral surfaces that the oil passes on its way from its source to the final accumulation (other authors^{2,3} have proposed that other compounds are also subject to differential adsorption). In consequence, the relative concentrations of the two isomers will change as the oil migrates further. This principle can potentially be used to tell how far the source rocks that gener-

ated the oils, and thus the petroleum systems that contain similar oils, extend geographically. It could thus be of considerable help in identifying viable new fields for exploitation.

This development has to be seen in the

per million years⁵), and is accompanied by changes in the petroleum composition. Secondary migration continues until barriers to flow are encountered. If the barrier rock is effectively a dead-end, the underlying reservoir rock constitutes a 'trap', and oil and/or gas will accumulate within it. Oil and gas explorers search diligently for such traps, most successfully by integrating geophysical, geological and geochemical data.

Molecular compounds used as migration tracers should ideally respond only to migration, and be as little affected as possible by bacterial degradation, differences in subsurface temperatures, and the type of organic matter in the source rocks. Most previously proposed oil-migration odometers^{6,7} are controversial, and probably little used, owing to their interdependence on these factors.

The benzocarbazoles that Larter and colleagues describe seem less likely to

have that flaw. They do not have alkyl side-chains, which, together with their structural configuration, should make them less liable to bacterial degradation. The lack of alkylation also means that they are less likely to interchange their structural configurations as subsurface temperatures increase. This means that changes of benzocarbazole ratios should be largely independent of how deeply buried the source rock was when it generated the several oils in a group of oils. These two compounds may, however, be present in quite different concentrations in source rocks that contain different types of organic matter. For example, a source rock with much marine-algal organic matter may yield low concentrations of benzocarbazoles, whereas source rocks in which woody-coaly organic matter predominates are more likely to yield higher concentrations.

It is also possible that an oil just starting secondary migration may exchange benzocarbazoles with those contained in the small amounts of organic debris in

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

A flare lights the night in the Urengoi gas field, West Siberia, the largest in the world. New exploration techniques and production technologies mean that fresh reserves of oil and gas can be identified and exploited even in well-worked fields such as this.

overall context of a 'petroleum system', which is a three-dimensional volume of the Earth's subsurface wherein oil and gas are generated, migrate upwards and laterally, and eventually are trapped below rock types that bar further flow. The potential source rocks, shales and carbonates, contain relatively high concentrations of organic matter, remnants of algal, bacterial and woody materials. When such rocks are buried deeply enough to raise temperatures into the range 60–200 °C, the organic matter is transformed to oil and gas, which are then expelled in a process known as primary migration. The proportions of each depend primarily on the chemical composition of the organic matter and the temperatures it experiences during burial; typically, oil is generated between 60 and 160 °C, gas 20 to 40 °C higher⁴.

The oil and gas then seek out secondary migration pathways that offer least pressures and higher permeabilities. On a geological timescale, the migration rate can be rapid (as much as 500–1,000 km

rocks such as sandstones. During a long secondary migration pathway, the oil would pass quite a bit of such debris. Originally, this debris may have contained concentrations and ratios of the benzocarbazoles that were quite different from those in oil just starting to migrate. If so, this would limit their use as migration odometers. Larter *et al.* may alleviate this problem by using ratios of the benzocarbazole concentrations in each group of related oils, and then normalizing the ratios to the ratio of the least-migrated (highest ratio) oil in the group. Further work may show that greater attention needs to be given to the type of organic matter in the source rock.

The key value of this new odometer lies in its use to better define the geographical extents of petroleum systems. If a source rock unit, having a similar organic-matter type throughout, generated oils at several geographically distant locations, then the ratios of the benzocarbazoles in several different oil samples may indicate how far the source interval, and thus the petroleum system, extends. These data would provide constraints for computer-based modelling of oil and gas generation from the source rock. The technique could also be applied to the identification of smaller reserves near existing petroleum accumulations already being exploited. Use of existing platforms and pipelines would then make such reserves economic. Also, detailed mapping of oil migration pathways into oil fields, as well as from fields to shallower traps, could help define where smaller, yet-to-be-found accumulations are located.

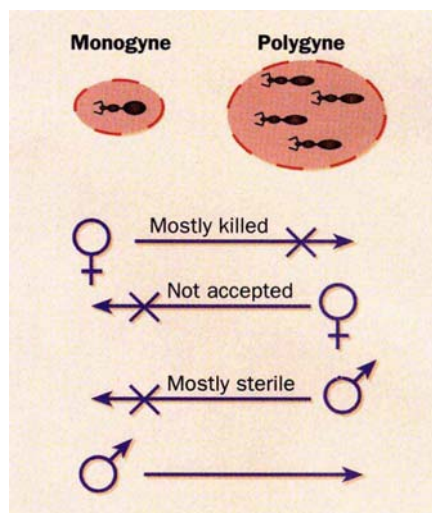
Many aspects of the secondary migration of petroleum remain poorly understood. These new migration odometers from Larter and colleagues look promising, however, and may provide answers to one of the main questions asked during exploration — “How far does the petroleum system extend?”. Petroleum geochemists will welcome them, and will be spurred to test them with field observations and experiments, and pursue the search for other such indicators. □

Gary H. Isaksen is in the Petroleum Geochemistry Section, Exxon Production Research Company, 3120 Buffalo Speedway, Houston, Texas 77252, USA (e-mail: gary.h.isaksen@eutec.exxon.com).

One into two will go

Ross H. Crozier and Pekka Pamilo

SPECIATION often includes a social dimension — after all, courtship is a social activity. But on page 613 of this issue¹, Shoemaker and Ross report a case of speciation in progress, one driven by basic changes to the overall social system of a eusocial insect, the imported fire ant. This is an example of sympatric speciation — that is, speciation within a geographically coexisting population (as opposed to allopatric speciation, which is well documented and occurs when a population becomes split by physical barriers).



Socially imposed speciation, three-quarters completed. Monogyne fire-ant nests have single maximally fertile queens, whereas polygyne nests have many queens of partially inhibited reproduction. Monogyne queens tend to have a genotype — as shown by homozygosity for two allozyme markers — that leads to their destruction in polygyne nests. Polygyne queens are killed if introduced into monogyne nests, and lack the fat reserves to found colonies independently in the monogyne manner. Most males produced by polygyne nests are sterile diploid males, so polygyne queens which mate outside the nest must usually mate with monogyne males, the only remaining route for gene flow. Polygyne queens also mate within their nests, and if this tendency increases, the final avenue for gene flow will disappear, and two species will have been created from one.

Introduced to the Southeastern United States from Argentina, the fire ant *Solenopsis invicta* has changed rapidly in its new environment. Initially, all of the ant colonies had a single queen (that is, were monogyne), and queens added experimentally were killed by the workers². Massive colonies with many unrelated queens then appeared (polygyne), with no close parallel in the Argentinian source area³. Although such colonies made life worse for those afflicted

with these ants, which have a sophisticated and unpleasant venom, they have provided a fascinating story of evolution in progress.

Queens in the monogyne form reproduce at apparently the maximum possible rate⁴, whereas those in the polygyne form have suppressed reproductive function (for example, smaller ovaries). Sex-determination in fire ants, as in many other Hymenoptera (ants, bees, wasps and sawflies)⁵, is probably controlled by a single locus at which heterozygotes are female and haploids and diploid homozygotes are male. But diploid males were not seen in fire ants until the advent of the polygyne form, because a single queen cannot survive the severe competitive stage in early colony life if half of her diploid progeny are male: queens in polygyne colonies do not face this form of selection because they join established colonies⁶. Unlike the mutual hostility shown between monogyne nests, polygyne nests are compatible, or ‘unicolonial’, and this increases their capacity for spread⁷.

Remarkably, there is a two-allele allozyme marker that correlates strongly with reproductive potential and is differentially selected between forms. Queens homozygous for the *Pgm-3^a* allele show heightened reproductive development compared with other genotypes. In polygyne, but not monogyne colonies, these queens are killed by workers when they start to mature after mating⁸. But the *Pgm-3^a* allele still occurs at appreciable frequencies in polygyne populations, despite the massive selection against it, because of gene flow from the monogyne population.

Four routes can be identified as potential conduits for gene exchange between the monogyne and polygyne forms, corresponding to possible migration by males and females of each form to the other (see figure). Shoemaker and Ross’s work¹ extends the nuclear-marker information to another differentially selected locus. Both forms of queen possess certain high-frequency mitochondrial DNA haplotypes (matrilineal markers) which are absent or rare in the other. This indicates that queens of one form do not move successfully into colonies of the other.

Both monogyne and polygyne forms leave the nest and mate during flight, but in the polygyne form an apparently increasing percentage of queens mate in the nest. Those polygyne queens that fly will usually mate with monogyne males, because most polygyne males are sterile diploids^{9,10}. If the tendency for intra-nest mating in polygyne colonies increases, gene flow

1. Larter, S. R. *et al.* *Nature* **383**, 593–597 (1996).

2. Snyder, L. R. *Principles of Adsorption Chromatography* (Dekker, New York, 1968).

3. Carlson, R. M. K. & Chamberlain, D. E. *Org. Geochem.* **10**, 163–180 (1986).

4. Hunt, J. M. *Petroleum Geochemistry and Geology* 2nd edn, 128 & 138 (Freeman, New York, 1996).

5. England, W. A., Mann, A. L. & Mann, D. M. *Treatise of Petroleum Geology* 23–46 (Am. Assoc. Petroleum Geologists, Tulsa, 1991).

6. Peters, K. E. & Moldovan, J. M. *The Biomarker Guide* 242 (Prentice-Hall, New York, 1993).

7. Krooss, B. M., Brothers, L. & Engel, M. H. in *Petroleum Migration* (eds England, W. A. & Fleet, J. A.) 149–163 (Geol. Soc., London, 1991).

will decrease and speciation will be complete. A corollary is that the various monogyne populations will probably stay genetically similar because of the extensive gene flow between them, but that the polygyne populations will become individually distinct because of their intraspecific mating.

Sympatric speciation has been proposed for ants before, but in a rather different way. Social parasites tend to be quite closely related to their hosts, and it has been suggested that they arise from intraspecific forms, specializing in entering established nests rather than founding colonies independently^{11,12}. The fire-ant speciation process may be applicable to ants generally, because there are many instances of sibling species where one is unicolonial and one is not. Crossing ants is difficult but not impossible, so that genetic dissection of this social evolution is achievable, and should help to decide between various hypotheses for speciation, such as a change in kin recognition cues or, more likely, a change in the receptivity to such cues¹³.

As far as we are aware, Shoemaker and Ross are the first to demonstrate that social behaviour influences sympatric speciation. But differences in social behaviour can facilitate gene flow as well as reduce it. One-way gene flow from anubis to hamadryas baboons occurs because hamadryas males sequester groups of females but anubis males do not — hamadryas males frequently begin female groups with kidnapped young anubis females¹⁴. Gene flow would cease if the social systems were the same — the opposite situation to the fire ants. □

Ross H. Crozier is at the School of Genetics and Human Variation, La Trobe University, Bundoora, Victoria 3083, Australia. Pekka Pamilo is in the Program of Conservation Biology, University of Uppsala, S-750 07 Uppsala, Sweden.

Master regulator unmasked

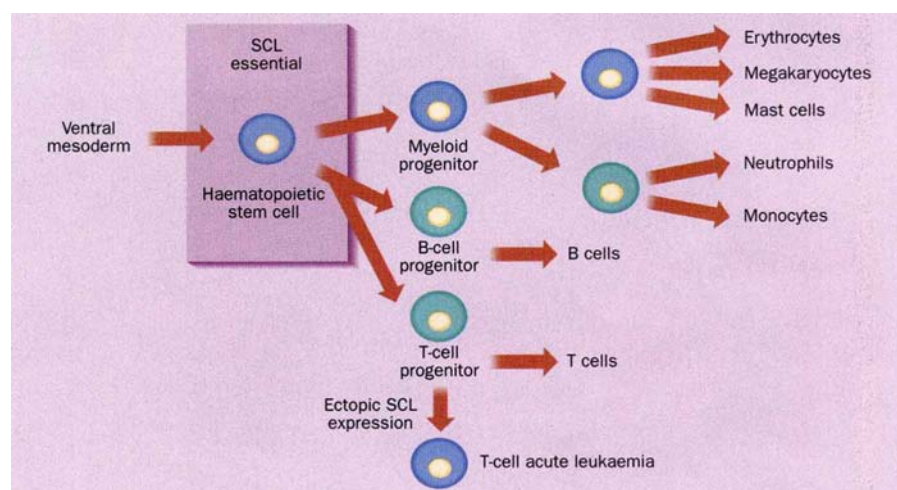
Tony Green

PLURIPOTENT stem cells are able to differentiate into several distinct cell types, but the mechanisms by which they do so remain one of the central mysteries of biology. Haematopoiesis (the production of blood cells) provides an excellent model for probing this process. In two recent papers by the Orkin¹ and Begley² groups, the product of the *SCL/TAL-1* gene (hereafter called *SCL*) is identified as a pivotal regulatory protein which is expressed in the pluripotent haematopoietic stem cell, and is essential for the development of all haematopoietic lineages.

The *SCL* (stem cell leukaemia) gene

SCL expression in committed haematopoietic lineages results from downregulation after the commitment of stem cells to non-expressing lineages, a process that may be necessary for normal differentiation along those lineages. *SCL* has since been found to inhibit apoptosis in leukaemic T cells⁴, whereas antisense and overexpression experiments indicate that *SCL* acts as a positive regulator of differentiation in erythroid cells, but regulates proliferation and self-renewal in multipotent cells^{5,6}.

These latter observations provided a strong hint that the knockout phenotype would be dramatic, and this has proved to



Two papers by the Orkin¹ and Begley² groups identify *SCL* as a pivotal regulatory protein which is essential during the early stages of blood cell (haematopoietic) development, and which also acts as a T-cell oncogene. Blue shading indicates cell types known or thought to express *SCL*. Ectopic *SCL* expression in a T-cell progenitor contributes to the development of T-cell acute lymphoblastic leukaemia.

encodes a basic helix-loop-helix transcription factor and, like many important developmental regulators, it was first identified by virtue of its rearrangement in a tumour (reviewed in ref. 3). Indeed, ectopic expression of *SCL* is perhaps the commonest molecular pathology associated with childhood T-cell acute lymphoblastic leukaemia. In murine systems, *SCL* increased tumour formation by a transformed T-cell line, and at least two different transgenic approaches have now demonstrated the oncogenic potential of *SCL* when expressed in T-cell progenitors.

The first clue to the normal function of *SCL* came from the characterization of its expression pattern. Apart from in the brain, *SCL* is expressed only in haematopoietic and endothelial cells, thus adding to the list of genes shared by these two cell systems. Within the haematopoietic system, *SCL* is expressed in primitive multipotent progenitors as well as in cells committed to the erythroid, mast and megakaryocytic lineages, but it is not found in several other cell types. So the pattern of

be so. Last year, the two groups^{7,8} reported that *SCL*-null mice died early in embryonic development and had no mature red blood cells. Neither erythroid nor myeloid progenitors could be detected in yolk sacs, the normal source of haematopoietic progenitors in the early embryo. These reports raised several questions. Was the absence of haematopoietic cells a cell-autonomous defect, or could the absence of *SCL* be affecting some critical aspect of the haematopoietic environment? Did the absence of myeloid progenitors merely reflect a secondary consequence of severe anaemia? Alternatively, perhaps *SCL* provided some essential function in pluripotent haematopoietic stem cells.

Fortunately, we have not had to wait long for the answers. The two groups^{1,2} have now described the generation of *SCL*-null embryonic stem cells, and their haematopoietic potential *in vitro* and *in vivo*. Their conclusions are the same — embryoid bodies derived from *SCL*-null embryonic stem cells do not contain any haematopoietic progenitors. Analysis of

- Shoemaker, D. D. & Ross, K. G. *Nature* **383**, 613–616 (1996).
- Wilson, E. O. in *Insect Behaviour* (ed. Haskell, P. T.) 81–96 (Royal Entomol. Soc., London, 1966).
- Ross, K. G., Vargo, E. L. & Keller, L. *Proc. Natl Acad. Sci. USA* **93**, 3021–3025 (1996).
- Ross, K. G. & Keller, L. *Ann. Rev. Ecol. Syst.* **26**, 631–656 (1995).
- Cook, J. M. *Heredity* **71**, 421–435 (1993).
- Ross, K. G. & Fletcher, D. J. C. *Evolution* **39**, 888–903 (1985).
- Bhatkar, A. P. & Vinson, S. B. in *Chemistry and Biology of Social Insects* (eds Eder, J. & Rembold, H.) 599–600 (Peppery, Munich, 1987).
- Keller, L. & Ross, K. G. *Science* **260**, 1107–1110 (1993).
- Ross, K. G., Vargo, E. L., Keller, L. & Trager, J. C. *Genetics* **135**, 843–854 (1993).
- Hung, A. C. F., Vinson, S. B. & Summerlin, J. W. *Ann. Entomol. Soc. Am.* **67**, 909–912 (1974).
- Buschinger, A. *Biol. Zbl.* **89**, 273–299 (1970).
- West-Eberhard, M. J. *Proc. Natl Acad. Sci. USA* **83**, 21388–21392 (1986).
- Crozier, R. H. & Pamilo, P. *Evolution of Social Insect Colonies. Sex Allocation and Kin-Selection* (Oxford Univ. Press, 1996).
- Gabow, S. A. *Am. Nat.* **109**, 701–712 (1975).

chimaeric adult mice (generated by injecting *SCL*-null embryonic stem cells into blastocysts), showed that the *SCL*-deficient cells could not give rise to any haematopoietic cell types, including T cells and B cells. Several other individual transcription factors have been shown to be essential for the development of multiple haematopoietic cell types⁹. But the *SCL* knockout is unique in that it completely abolishes all lineages and their precursors. The *SCL* protein therefore provides a critical cell-autonomous activity that is essential for the development or behaviour of pluripotent haematopoietic stem cells (see figure). This is reminiscent of the requirement for other basic helix-loop-helix factors during the development of muscle or brain.

How does *SCL* perform this pivotal function? Here we must return to studies designed to elucidate how *SCL* acts as a T-cell oncogene. It had been clear for some time that *SCL* binds to DNA as a heterodimer with widely expressed class A helix-loop-helix transcription factors¹⁰. More surprising was the subsequent demonstration of a direct interaction between *SCL* and a second T-cell oncoprotein, the LIM-domain protein LMO-2 (previously known as RBTN-2 or TTG-2)^{11,12}. Although this observation was unexpected because it was the first example of an interaction between a LIM-domain protein and a helix-loop-helix transcription factor, it was consistent with the concept of oncogene collaboration in tumorigenesis. But this interaction may also be relevant to the normal function of *SCL* because LMO-2 is coexpressed with *SCL* in normal haematopoietic cells, and its knockout phenotype is remarkably similar to that of *SCL*¹³. So these results indicate that a complex containing *SCL* and LMO-2 may be involved not only in some T-cell tumours, but also in normal haematopoietic cells.

The story doesn't finish there. In several haematopoietic cell types, *SCL* is coexpressed with two other key regulators of haematopoiesis, GATA-1 and GATA-2, both of which are transcription factors containing zinc-finger motifs. Moreover, LMO-2 can bind to both GATA-1 and GATA-2 as well as to *SCL*. Therefore, although LMO-2 seems not to bind to DNA itself, it may act as a physical bridge between *SCL* and the GATA proteins to form a tetrameric com-

plex¹⁴. The physiological significance of this complex will need to be confirmed because the amount of LMO-2 complexed with GATA-1 in erythroid cells is very low relative to the amount of free, unbound GATA-1. But the concept of a tetrameric complex is thought-provoking, not least because it implies that there are several layers of complexity. First, the concentrations of the different GATA proteins would vary between different haematopoietic cell types. For example, GATA-1 predominates in erythroid cells whereas GATA-2 is more abundant in multipotent progenitors. Second, *SCL* can interact with several different widely expressed helix-loop-helix proteins. Third, *SCL* is regulated by GATA-1 in erythroid cells and so at least some components of the tetrameric complex can regulate one another. And fourth, further consternation for the faint-hearted is provided by the intricate network of equilibria likely to exist between monomers, homodimers, heterodimers and the putative tetrameric complex.

ARCHAEOLOGY

Further back down under

Paul G. Bahn

EVIDENCE of rock engravings in Australia dating back at least 75,000 years, and of stone tools up to about 176,000 years old, will be published in December in the journal *Antiquity*¹. This news has greatly exercised the world's press. If true, both discoveries will have profound consequences for the study of human origins and world prehistory.

A team of three scientists led by Richard Fullagar made the discoveries a few years ago, at a remote site called Jimmim in the Northern Territory, near the Ord River. Here, since 1992, stone tools have been found in a layer of sediment dated to about 60,000 years ago. There are stone chips and flakes, as well as pounding tools that apparently bear microscopic traces of starchy residues, implying that they were used for processing plant material, probably tubers.

The site contains enormous sandstone boulders, some of which are covered with thousands of small, pecked cupmarks — 3,500 have been counted on the rock face immediately above excavated areas, and 3,200 on another nearby. A detached fragment bearing these pits lies in sediments dated by the thermoluminescence technique to between 58,000 and 75,000 years ago. Lower layers of sediment nearby, containing ochre and stone artefacts, have likewise been dated to 116,000 ± 12,000 years ago, and even 176,000 ± 16,000 years for a layer just below one object.

The marks on the rocks are definitely artificial, and it has been estimated that

A whole host of questions now loom large. *SCL* and LMO-2 are thought to be coexpressed in only a minority of cases of T-cell acute lymphoblastic leukaemia. What is happening in tumours that express only one of these proteins? Can other proteins interact with *SCL* or LMO-2 in these tumours, or are completely different mechanisms operating? It remains possible that, in at least some cases, *SCL* may not directly activate or repress target genes, but it may function in a dominant-negative manner to sequester other transcriptional proteins. With respect to the normal function of *SCL*, it will be crucial to identify the target genes that it regulates in normal haematopoietic stem cells and in lineage-committed cells. And finally, what are the mechanisms that determine the pattern of *SCL* expression during haematopoiesis — in other words, who regulates the regulator? □

Tony Green is in the Department of Haematology, University of Cambridge, MRC Centre, Hills Road, Cambridge CB2 2QH, UK.

each cupule would have taken at least an hour to make. That fact underlines their importance to the people who produced them, as the thousands found so far would have taken one person almost 900 eight-hour days to complete (so they were probably produced by a larger number of people or accumulated over a long period).

The authors need to persuade other archaeologists that the very early objects are indeed artefacts. One enigmatic circular stone from a layer between 75,000 and 116,000 years old is unlike anything seen before in a prehistoric or modern Aboriginal tool kit, and lacks microscopic traces of use (wear, polish or residues), but they believe that the rounded outer edges are evidence for its having being worked.

But the greatest scepticism is over the dates and the technique used to obtain them. Radiocarbon and accelerator mass spectrometry techniques normally have a limit of 40,000 to 50,000 years ago, and need organic material, so they are of little use in verifying these results. Luminescence techniques, on the other hand, have a potential range of 30,000 to 300,000 years, and require no organic material. They measure the time that grains of quartz have been buried and hidden from sunlight, by counting the electrons (produced by background radioactivity) that get trapped in microfissures in the grains — these electrons are rapidly lost on exposure to sunlight, but accumulate in grains that are buried. If a sample of sand is heated or exposed to laser light, the

1. Porcher, C. et al. *Cell* **86**, 47–57 (1996).
2. Robb, L. et al. *EMBO J.* **15**, 4123–4129 (1996).
3. Robb, L. & Begley, C. G. *Int. J. Biochem. Cell Biol.* **28**, 609–618 (1996).
4. Leroy-Viard, K. et al. *EMBO J.* **14**, 2341–2349 (1995).
5. Green, A. R. et al. *EMBO J.* **10**, 4153–4158 (1991).
6. Aplán, P. D. et al. *EMBO J.* **11**, 4073–4081 (1992).
7. Robb, L. et al. *Proc. Natl Acad. Sci. USA* **92**, 7075–7079 (1995).
8. Shivdasani, R. A. et al. *Nature* **373**, 432–434 (1995).
9. Shivdasani, R. A. & Orkin, S. H. *Blood* **87**, 4025–4039 (1996).
10. Hsu, H. L. et al. *Mol. Cell Biol.* **11**, 3037–3042 (1991).
11. Valge-Archer, V. E. et al. *Proc. Natl Acad. Sci. USA* **91**, 8617–8621 (1994).
12. Wadman, I. et al. *EMBO J.* **13**, 4831–4839 (1994).
13. Warren, A. J. et al. *Cell* **78**, 45–57 (1994).
14. Osada, H. et al. *Proc. Natl Acad. Sci. USA* **92**, 9585–9589 (1995).

released electrons produce an amount of light that is determined by the duration of the burial.

The two principal luminescence techniques are thermoluminescence (TL) and optically stimulated luminescence (OSL). Both have already been used in Australia to date early archaeological sites, indicating that other sites in the Northern Territory were occupied up to 60,000 years ago, for example². But these dates are controversial, and there are problems with the technique — in particular, the sample of sand may contain grains that have been hidden from sunlight for millions of years, from rockfalls or decomposed bedrock that has become mixed into the sample through wind or water action. Such intrusive grains can produce ages that are far too great. The researchers acknowledge that this is a difficulty that may affect the lowest level, with its 176,000-year date.

There is, however, a new technique, the OSL dating of individual sand grains, that can root out anomalously old readings. It has already been used to rectify some very old (unpublished) TL dates for sediments in a Queensland occupation site, and — in view of the abundance of rockfall in the Jinnium site — must certainly be applied to the new claims. The team are attempting to have this done.

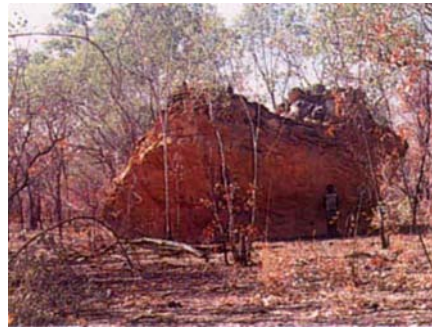
If the Jinnium dates prove to be valid, what are the implications for rock art and prehistory? They are little short of revolutionary for those few scholars who cling to the traditional assumptions that 'art' began with modern humans and the Upper Palaeolithic period (from about 40,000 years ago) and that there was no human occupation of Australia before 60,000 years ago. Yet these assumptions have been, respectively, demolished and seriously called into question in recent years, so the Jinnium findings reinforce the convictions of those who have espoused a less traditional approach to the past.

Contrary to the claims in the press, the oldest known rock art in the world is not that of the French cave of Chauvet, where charcoal in the paint has produced radiocarbon ages of up to 32,000 years. There are older examples elsewhere. In Australia, at the rock shelter of Carpenter's Gap, what seems to be a fallen fragment of painted wall was found in a layer containing charcoal more than 39,000 years old³; and the (albeit controversial) dating of rock varnish covering petroglyphs in South Australia⁴ gave ages up to more than 43,000 years.

Even older art, attributed to *Homo erectus*, has been discovered. Israeli archaeologists have found a small piece of volcanic tuff, between 233,000 and 800,000 years old, at Berekhat Ram on the Golan Heights⁵. It resembles a female figurine, with grooves around its neck and along its arms. Much rested on the question of whether these grooves were natural or arti-

ficial, but microscopic analysis has now shown clearly that human hands were responsible. This was an intentionally enhanced image, indisputably an 'art object'.

Other evidence has appeared from the same period, such as bones from the open-air site of Bilzingsleben, Germany, which bear a series of what seem to be decorative parallel incisions⁶. Where rock art is concerned, two petroglyphs — a large circular cupule and a pecked



Property with original features: this Australian site may have been occupied 176,000 years ago. (Photo courtesy of *Antiquity*.)

meandering line — have been found in Auditorium Cave at Bhimbetka, India, covered by an acheulian occupation layer⁷ (probably several hundred thousand years old). If both this example and Jinnium are valid, they will confirm the suspicions of many scholars that cupules and other simple motifs are the oldest surviving rock markings in many parts of the world. Traditional theories that art began in the Upper Palaeolithic, or with modern humans, let alone in Europe or even the Dordogne, are now utterly discredited.

As for the colonization of Australia, it was only in the 1960s that firm evidence was found for human occupation during the Pleistocene, before about 10,000 years ago; since then, the date has been steadily pushed back, always in the face of enormous scepticism. The currently accepted arrival date is up to 60,000 years ago⁸, yet over the past decade there have been strong claims from pollen studies for a far earlier date. Sediment cores from Lake George, New South Wales, revealed a period around 130,000 years ago when there was a sudden increase in destructive bush fires, reflected in greatly increased quantities of charcoal, and coinciding with a sudden and dramatic change in vegetation — the first for 750,000 years — when fire-sensitive forests began to be replaced by fire-tolerant eucalyptus⁸ and by grasses.

That was challenged by archaeologists⁹, who assigned this vegetation change to only 60,000–54,000 years ago, using a simple correlation of age with depth in the lake cores. But support for the early dates has come from similar pollen and charcoal evidence, dated to about 140,000 years ago, in a 400-metre marine core on the continental shelf off the Queensland coast¹⁰; the core spans 1.5 million years,

and, unlike Lake George, its upper part is well dated. The decline of forest and rise of charcoal particles form the most dramatic change in the whole core, and have been tentatively attributed to human colonization. Once again, however, there have been vociferous challenges from archaeologists and others^{11–14}.

One of the stumbling blocks to accepting an early colonization has always been that, because Australia was never attached to Asia, even at the lowest sea levels of the Ice Age a journey there required an ocean voyage of at least 70 km. But now the discovery of a stone-tool industry on the Indonesian island of Flores, in a layer dated by palaeomagnetism to about 700,000 years ago, has provided good evidence for open-sea voyages by *H. erectus*¹⁵, so there can be no question that whoever colonized Australia was capable of such a journey, even 176,000 years ago.

The Jinnium claims may also have tremendous repercussions for the debate between the two main models of human origins. According to the 'Out of Africa' theory, modern humans arose in Africa and migrated from there around 100,000 years ago — if that is so, then the Jinnium people, should their date be validated, cannot have been modern humans. But those who support the hypothesis of 'multi-regional continuity' argue that early humans left Africa about 2 million years ago, and have been evolving and spreading elsewhere since then, intermixing enough to explain the parallel evolution. Some have already speculated that the Jinnium people were some kind of late survival of an archaic form of *Homo sapiens* or even that they were *H. erectus*, and that they evolved in Australia into modern hominid forms. But it is far too early to reach any firm conclusions about the identity of the colonizers.

A great deal hangs on the corroboration of the Jinnium claims. It is to be hoped that OSL checks will eventually reveal whether the media attention was justified, and whether so many of the established ideas about the remote past now need to be jettisoned. □

Paul G. Bahn writes from 428 Anlaby Road, Hull HU3 6QP, UK.

- Fullagar, R. L. K., Prince, D. M. & Head, L. M. *Antiquity* (in the press).
- Roberts, R. G. et al. *Quat. Geochronol.* **13**, 575–583 (1994).
- O'Connor, S. *Aust. Archaeol.* **40**, 58–59 (1995).
- Nobbs, M. F. & Dorn, R. I. *Archaeol. Oceania* **28** (1), 18–39 (1993).
- Goren-Inbar, N. *Mi'tekufat Ha'even* **19**, 7–12 (1986).
- Bednarik, R. G. *Curr. Anthropol.* **36**, 605–634 (1995).
- Bednarik, R. G. *Man Envir.* **18** (2), 33–40 (1993).
- Singh, G. & Geissler, E. A. *Phil. Trans. R. Soc. Lond. B* **311**, 379–447 (1985).
- Wright, R. *Archaeol. Oceania* **21**, 138–139 (1986).
- Kershaw, A. P. *The Artefact* **16**, 12–18 (1993).
- White, J. P. *Quat. Aust.* **12** (2), 21–23 (1994).
- Kershaw, A. T. *Quat. Aust.* **12** (2), 23–29 (1994).
- Anderson, A. *Quat. Aust.* **12** (2), 30–31 (1994).
- Hope, G. *Quat. Aust.* **12** (2), 32–33 (1994).
- Sondaar, P. Y. et al. *C. R. Acad. Sci.* **319**, 1255–1262 (1994).

A little help for my ends

Jef D. Boeke

REVERSE transcriptases (RTs) are usually perceived as engines of devastation employed by evil retroviruses such as HIV. But RTs also form an essential component of the replication machinery of a dizzying array of elements, including DNA viruses, plasmids, mobile introns and diverse retrotransposons. The latter are an important group of genome parasites that replicate, much like retroviruses, through reverse transcription mechanisms, thereby increasing their copy number in the host genome. Although the potential deleterious effects of these elements on their host species are obvious (occasional host gene inactivation and increased DNA burden), there are hints that retrotransposons may occasionally benefit their hosts.

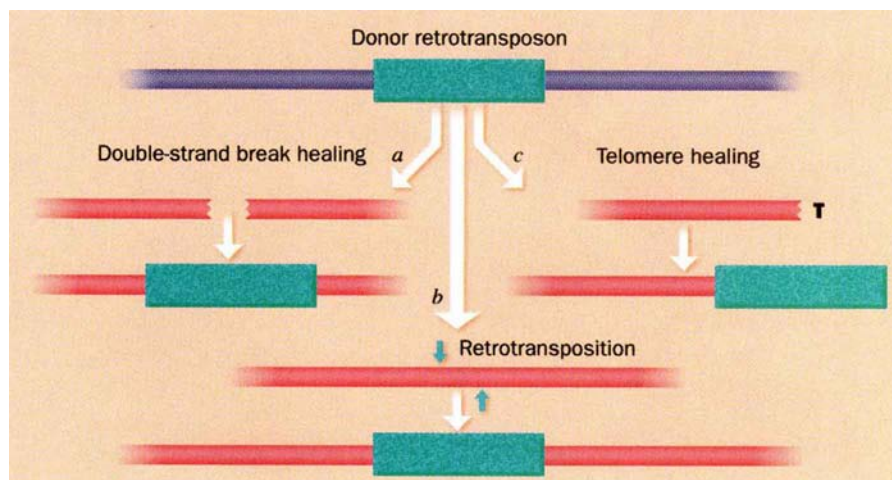
One such indication now comes from the Gabriel¹ and Haber² laboratories. On pages 641 and 644 of this issue, respectively, they show that retrotransposons and the RTs they encode can function as 'molecular band-aids' to heal potentially lethal chromosome breaks. Both of these groups studied the repair of double-strand breaks that had been generated by over-expressed HO endonuclease in *Saccharomyces cerevisiae*. These breaks are lethal in yeast strains lacking the *RAD52* gene, because homologous repair (use of a homologous DNA strand as a template) is genetically blocked.

Teng, Kim and Gabriel¹ describe the analysis of a large number of rare survivors of HO cleavage. They exploited a pseudogene capture assay to show that RTs from three different retrotransposons can effect break repair. They then used this system to select for large numbers of DNA break-repair events mediated by RTs, and found that Byzantine patchworks of pseudogene and retroelement sequences were inserted at the break sites.

Using a similar HO/*rad52*-survival strategy, Moore and Haber² found that about one per cent of the survivors contained a ~100-base-pair fragment of the endogenous yeast retrotransposon Ty1, which had been inserted at the break site. Remarkably, four out of five of these insertions corresponded to the nearly precise acquisition of an important Ty1 replication intermediate called minus-strand strong stop DNA at the endonuclease cleavage site. The experiment was also performed in recombination-proficient *RAD52* strains in which homologous repair was blocked by other means, and similar insertions were observed.

In both studies, DNA fragments bearing the stigmata of reverse transcription processes were inserted at the break site (a in the figure). The mechanisms used by

retrotransposons to prime DNA synthesis are diverse. They include priming by a cellular transfer RNA in many elements such as Ty1, which have long terminal repeats (LTRs)³; self-priming by a small RNA fragment⁴; priming by an internal 2' OH in the RNA of prokaryotic retrons⁵; and priming from nicks in target DNA, as in the non-LTR retrotransposons⁶. Such priming events normally give rise to full-length retrotransposon DNA, which becomes inserted as an intact element into a new target site (b in the figure).



Alternative fates of retrotransposon sequences. a, b, Teng *et al.*¹ and Moore and Haber show that partial retrotransposon sequences can heal intrachromosomal double-strand breaks. The mechanism of this reaction is not known. The major fate of donor retrotransposon sequences (blue) is the formation of new intact retrotransposon copies through retrotransposition into intact chromosomal DNA targets (pink). Blue arrows indicate breaks made by retrotransposon-encoded integrases. c, Certain specialized retrotransposons, such as HeT-A and TART, can apparently serve as new telomeres for broken telomeres (T). This probably occurs by a process resembling retrotransposition^{6,9}.

Finally, at least one retroelement RT, from the Mauriceville mitochondrial plasmid of *Neurospora*, can carry out primer-independent synthesis⁷. Thus, it should not be so surprising that this adaptable enzyme can be pressed into service to patch up breaks in chromosomal DNA.

One particularly relevant form of priming by RT has been proposed for the TART element, a non-LTR retrotransposon of *Drosophila*. Whereas most eukaryotic chromosomes have short telomerase-generated nucleotide sequence repeats at their termini, *Drosophila* termini comprise copies of the TART element, together with a related element, HeT-A. These elements are able to heal broken chromosome ends in *Drosophila*, so halting the inexorable shortening of 50–100 base pairs per generation owing to incomplete terminal replication⁸. Levis⁹ has proposed that TART RT can prime reverse transcription using the 3' end of a broken chromosome as a primer and TART RNA as a template (c in the figure).

The structural analysis of 'healed' chromosomes as well as native *Drosophila* telomeres supports just such a model. Notably, certain LTR retrotransposons are also common in or near telomeres, and at least one of these — Ty5 — actively targets telomeric DNA for new insertions¹⁰. The studies by Teng *et al.* and by Moore and Haber point towards possible mechanistic connections to the TART telomere-formation model. They also indicate that the ability to repair breaks, as part of an 'emergency break-repair system' for the cell, may be evolutionarily conserved among retrotransposons.

In yeast cells, the vast majority of double-strand breaks are healed by the general homologous-recombination sys-

tem, which relies on the *RAD52* gene product. But when this pathway is inactivated by mutation, or cannot operate because a homologous segment of DNA that can serve as the donor of healing information is unavailable, backup repair processes come into operation. These allow breaks to be healed by the addition or deletion of a few base pairs at the break

1. Teng, S.-C., Kim, B. & Gabriel, A. *Nature* **383**, 641–644 (1996).
2. Moore, J. K. & Haber, J. E. *Nature* **383**, 644–646 (1996).
3. Chapman, K. B., Bystrom, A. S. & Boeke, J. D. *Proc. Natl. Acad. Sci. USA* **89**, 3236–3240 (1992).
4. Levin, H. L. *Mol. Cell. Biol.* **16**, 5645–5654 (1996).
5. Shimamoto, T., Inouye, M. & Inouye, S. *J. Biol. Chem.* **270**, 581–588 (1995).
6. Luan, D. D., Korman, M. K., Jakubczak, J. L. & Eickbush, T. H. *Cell* **72**, 595–605 (1993).
7. Wang, H. & Lambowitz, A. M. *Cell* **75**, 1071–1081 (1993).
8. Biessmann, H. *et al. Cell* **61**, 663–673 (1990).
9. Levis, R. W., Ganesan, R., Houchens, K., Tolar, L. A. & Sheen, F.-M. *Cell* **75**, 1083–1093 (1993).
10. Zou, S., Ke, N., Kim, J. M. & Voytas, D. F. *Genes Dev.* **10**, 634–645 (1996).
11. Moore, J. K. & Haber, J. E. *Mol. Cell. Biol.* **16**, 2164–2173 (1996).

site. These healing events generally resemble the ill-defined pathways of non-homologous recombination that mediate the integration of transfected DNA into mammalian chromosomes¹¹.

The presence of the products of reverse transcription precisely at the break site raises many questions about the mechanism by which the reverse transcripts heal the break. Is this an active mechanism in which a specific priming event by RT occurs at the break site? Or is it more passive, with previously synthesized patches of reverse transcripts being cobbled into the break by the non-homologous-recombination machinery? In the case of the Ty1 RT-mediated healing, a specific reverse-transcription intermediate, minus-strand strong stop DNA, which is covalently associated with a transfer RNA at its 5' terminus during Ty1 reverse transcription, seems to be essential. Yet paradoxically,

overexpression of a wild-type Ty1 element fails to stimulate this type of healing². Perhaps mutant Ty1 elements present in the genome donate this DNA intermediate. In particular, protease mutants, with their unstable capsids, might be a likely source of this molecule.

Whatever the detailed mechanism, it seems likely that the HO enzyme remains associated with the newly cleaved target DNA ends for some time, preventing the broken ends from drifting apart inside the cell before the repair machinery can get hold of them. Do RTs have a primordial inherent affinity for such nicked complexes that might allow them to help out a cell in need? Or do they respond to a regulatory system that detects DNA breaks, signalling a need for various repair activities? The genetic systems developed by Teng *et al.*¹ provide a useful tool kit for exploring the mechanism and regulation

of these retrotransposon-mediated repair events.

These new findings hint at the existence of hitherto unknown mechanisms for DNA rearrangements that can shape genomes. Eukaryotic chromosomes are a rich tapestry of retrotransposons and other repeat sequences, often disparaged as junk DNA. Processed pseudogenes are common in human and other mammalian cells, as are Alu sequences and related retrotranscripts. Perhaps some of these insertions represent ancient patchings of breaks in our ancestors' chromosomes, and we owe our current linkage map, at least in part, to some of the helpful aspects of our DNA parasites. □

Jef D. Boeke is in the Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205-2185, USA.

COMETS

Hyakutake's interstellar ices

Michael J. Mumma

COMETS are messengers from the birth of our planetary system — so what can they tell us about the process leading from the original interstellar cloud to our system of planets? Particularly, how much were ices from the natal cloud modified in the pre-planetary nebula before being incorporated into cometary nuclei¹⁻³? The answer is of interest because much of Earth's water and its early organic chemicals were delivered by comet impact, after the primordial atmosphere had been lost^{4,5}. The chemistry of interstellar cloud cores^{6,7} leads to quite different abundances of organic compounds from that of the warmer, denser pre-planetary nebula^{2,8}. So to assess the significance of cometary organic compounds in the evolution of the Solar System and the origins of life, the first step is to ask, are they of interstellar origin? New observations of the chemicals given off by the comet Hyakutake provide a partial answer: yes, in some cases.

On page 606 of this issue⁹, Brooke *et al.* report the detection of acetylene — a gas commonly used for welding metals on Earth — in an abundance consistent with that predicted for ices in interstellar dense clouds. And two weeks ago, Irvine *et al.*¹⁰ reported that the abundance ratio for two forms of hydrogen cyanide in Hyakutake is consistent with the value measured for interstellar

cloud cores. Taken with reports of abundant ethane and methane in this comet¹¹, these findings suggest strongly that Hyakutake contains ice that formed in our natal interstellar cloud.

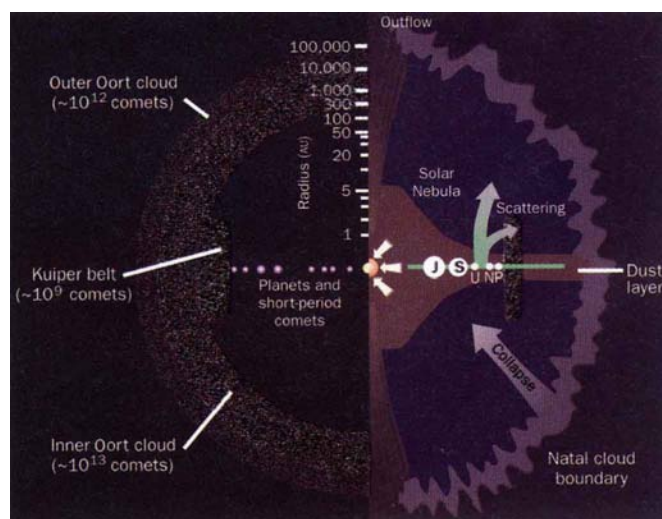
It is generally agreed that a star and its planetary system form by the collapse of a dense cloud of interstellar dust and gas¹² (see figure). Temperatures in these clouds are so low that most common gases condense onto dust grains as ice, and only gases with ultra-low vaporization temperatures — such as hydrogen, helium and

nitrogen — remain in the vapour phase. But as the natal cloud collapses, its temperature increases. Dust grains in distant regions of the proto-planetary disk may retain their icy mantles¹, but those close to the proto-Sun lose them, and the vaporized gases may react chemically^{2,8} before condensing again as the nebula cools. Comets are thought to have formed by the accumulation of icy grains beyond Jupiter¹³, and to have remained largely unchanged since then, so their compositions should reveal the origin of cometary ices³.

But is Hyakutake a typical comet? Comets today come from two principal reservoirs. The Edgeworth-Kuiper disk ranges from Neptune, at about 30 astronomical units (AU) to perhaps a few hundred AU from the Sun. The

best known member of this comet family is Pluto, and several dozen other large ones (more than 100 km in diameter) have been found since 1992. These bodies have remained in this region since being formed, but occasionally one is gravitationally perturbed onto an orbit that carries it into the inner Solar System, where it is activated by sunlight and seen as a short-period comet.

The second reservoir is called the Oort cloud, after Jan Oort, who first proposed its existence (but this is also an unintended word-play on the orts, or leavings, of the Solar System). These comets formed closer to the Sun — between 5 and 30 AU — and were ejected to very great distances by gravitational interactions with the giant planets. Originally a flat-



Before and after. On the right, a cloud of gas collapses to form our Solar System. After about ten million years, the Sun is forming within the Solar Nebula, and comets fill the early Kuiper belt; others, further in, get scattered by the young giant planets. Within another million years (left), the present-day distribution of planets and comets is already established. Ices in the comet Hyakutake appear to have come from the original interstellar cloud.

tened disk of comets with orbital inclinations close to those of the giant planets, the galactic tide has randomized their orbital inclinations and raised their perihelia out of the giant-planet region, creating a spherical cloud of about 10^{13} comets at distances of order 10,000 AU (ref. 14).

Passing stars occasionally thread the Oort cloud, and perturb some comets into the inner Solar System, where they are seen as 'dynamically new' comets. Their gravitational interaction on this first passage, mainly with Jupiter, often perturbs them into much smaller orbits. Hyakutake is one such comet, as is Halley, and the coming super-comet Hale-Bopp is yet another.

If ices in the Jupiter-to-Neptune zone of the collapsing nebula have a common physical and chemical history, the compositions of all these long-period comets should be the same. But that is clearly not the case; the abundance of the various ices (methanol, for example) varies significantly among members of this population. Hyakutake had a much higher ethane abundance than can be accommodated by infrared signatures of other comets, and neither acetylene nor ethane was detected in Halley. If comets are pristine, these differences imply a radial or temporal gradient in the composition of pre-cometary ices in the 5–30-AU region of the nebula. Current theories do not consider such gradients, and so may need to be revised.

Interstellar acetylene ice has never been detected, but the predicted abundance in dense molecular-cloud cores⁶ is comparable to that measured by Brooke *et al.*⁹ in Hyakutake, and both acetylene (C_2H_2) and carbon monoxide can be produced in the gas phase in interstellar clouds. Ethane (C_2H_6) cannot, but it can be made by adding hydrogen atoms to acetylene condensed onto grain surfaces^{6,7,15}. The high ratios of ethane to

acetylene in Hyakutake are consistent with this process¹¹, and with the accompanying hydrogenation of CO to H_2CO . The low abundance of methanol (which would be similarly hydrogenated, from formaldehyde) then requires a low formation temperature¹⁶, perhaps less than 30 K.

This temperature is supported by the ortho/para ratio for H_2O and the D/H ratio of the water in Halley, both of which suggest the formation of cometary water at about 30 K in the natal interstellar cloud³. And the high methane abundance in Hyakutake¹¹ is consistent with condensation fractionation from the gas phase at these temperatures.

The isomeric abundances are also consistent with an interstellar origin at low temperatures. Irvine *et al.*¹⁰ argue that the high HNC/HCN ratio in Hyakutake is consistent with values found for cold, dense interstellar cloud cores, where chemical enrichment by temperature-sensitive ion-molecule reactions can lead to a ratio of 1 near 10 K, decreasing to about 0.005 at

200 K. It is possible, however, that the HCN gets modified in the coma, chemically or by solar radiation, to yield the observed large enrichment in HNC/HCN in Hyakutake. But the HCN abundance alone suggests an interstellar origin; it is too high to be consistent with gas-phase processing in the proto-planetary disk².

How, then, should we interpret the Hyakutake results? Was the comet formed near Neptune from interstellar ices? Was it formed closer to the proto-Sun from ices processed in the nebula? Do other comets resemble Hyakutake in their chemistry? We must probe more deeply into the chemistry of comets to fully comprehend their role as cosmogonic messengers, and their part in delivering the pre-biotic organic compounds needed for the origin of life. □

Michael J. Mumma is at the Laboratory for Extraterrestrial Physics, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA.

DNA REPLICATION

Tussle with a terminator

R. G. Wake

TOWARDS the end of a round of replication of a circular bacterial chromosome, two approaching replication forks are forced to meet within a restricted 'terminus' region, approximately opposite the origin. This is achieved through the presence of a replication-fork arrest system in this region (Fig. 1). Such systems were first discovered, and have been most

extensively characterized, in *Escherichia coli* and *Bacillus subtilis*, and it is likely that they allow the more efficient coupling of chromosome replication and cell division¹. The two bacterial systems each comprise several short DNA terminators (*Ter*) of 20–30 base pairs, plus a cognate terminator protein that recognizes *Ter* sites and binds to them tightly².

On page 598 of this issue, Kamada and colleagues³ report the X-ray crystal structure of the *E. coli* terminator protein (Tus)–*Ter* complex. This is the first structure of a terminator protein–*Ter* complex to be elucidated, and it is a major advance — the way in which Tus fixes itself to the DNA is completely new amongst DNA-binding proteins. Furthermore, the authors have devised a persuasive model for how the unique structure, by itself, can explain the polarity of replication-fork arrest.

The most intriguing aspect of a terminator protein–*Ter* complex is its polarity: the ability to arrest a fork approaching from one direction but not from the other. The arrest complex is thought to work by blocking the helicase-mediated unwinding of duplex DNA at the apex of the fork. But there are differing views as to how this might be achieved. Some argue that the protein–DNA complex acts as a general barrier to the progression of the helicase⁴. Others suggest that specific interactions between the terminator protein and the helicase are involved^{5,6}.

Previous work established that Tus binds

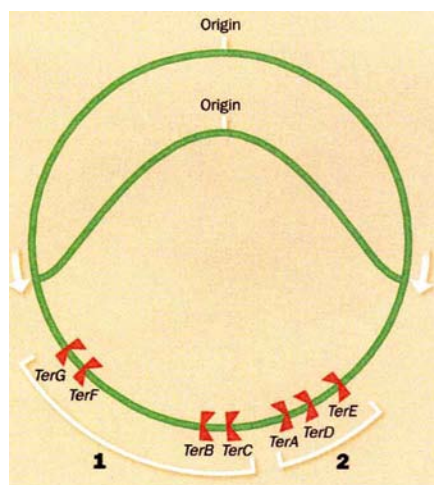


FIG. 1 A partially replicated *E. coli* chromosome with double-stranded DNA shown as a single line. Two replication forks (arrows), initiated at the origin, approach each other and encounter two groups of opposed terminators (*Ter* sites). Those in group 1 can arrest progression of the clockwise fork, whereas those in group 2 are orientated to block the anticlockwise fork. This ensures that the forks always meet within the region spanned by the terminators.

- Greenberg, J. M. in *Comets* (ed. Wilkening, L. L.) 131–163 (Univ. Arizona Press, Tucson, 1982).
- Fegley, B. Jr & Prinn, R. G. in *The Formation and Evolution of Planetary Systems* (eds Weaver, H. A. & Danly, L.) 171–205 (Cambridge Univ. Press, 1989).
- Mumma, M. J., Weissman, P. R. & Stern, S. A. in *Protostars and Planets Vol. III* (eds Levy, E. H., Lunine, J. I. & Matthews, M. S.) 1172–1252 (Univ. Arizona Press, Tucson, 1993).
- Chyba, C. F. *Nature* **343**, 129–133 (1990).
- Chyba, C. & Sagan, C. *Nature* **355**, 125–132 (1992).
- Hasegawa, T. I. & Herbst, E. *Mon. Not. R. Astron. Soc.* **263**, 589–606 (1993).
- Charnley, S. B., Kress, M. E., Tielsens, A. G. G. & Millar, T. J. *Astrophys. J.* **448**, 232–239 (1995).
- Eck, R. V., Lippincott, E. R., Dayhoff, M. O. & Pratt, Y. T. *Science* **153**, 628–633 (1966).
- Brooke, T. Y. *et al.* *Nature* **383**, 606–608 (1996).
- Irvine, W. M. *et al.* *Nature* **383**, 418–419 (1996).
- Mumma, M. J. *et al.* *Science* **272**, 1310–1314 (1996).
- Beckwith, S. V. W. & Sargent, A. I. *Nature* **383**, 139–144 (1996).
- Weidenschilling, S. *Nature* **368**, 721–723 (1994).
- Duncan, M., Quinn, T. & Tremaine, S. *Astrophys. J.* **328**, L69–L73 (1988).
- Tielsens, A. G. G. & Hagen, W. *Astron. Astrophys.* **114**, 245–260 (1982).
- Hiraoka, K. *et al.* *Chem. Phys. Lett.* **229**, 408–414 (1994).

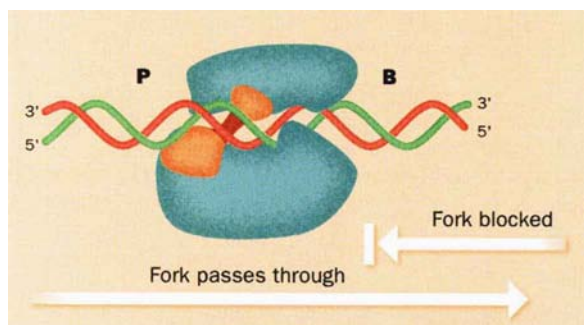


FIG. 2 Model of the Tus (protein)-Ter (DNA) complex, devised by Kamada *et al.*³, in which the DNA is shown to extend beyond the contact region at each end. A replication fork moving from right to left (arrow) is arrested at the B (blocking) end. But a fork moving in the opposite direction is not impeded upon encountering the P (passage) end and passes through the complex freely.

to each Ter site as a monomer^{7,8}. Within the Tus-Ter complex (Fig. 2) the protein is organized into two domains, each composed of α -helical (blue) and β -sheet (yellow) portions, connected by two antiparallel β -strands (brown). This creates a large, positively charged central cleft into which double-stranded DNA fits, such that the DNA sits across the interdomain region, and is flanked by the two large protein domains. The DNA deviates significantly from the Watson-Crick B-form — the DNA is underwound, and the backbone is deformed in regions where the DNA makes particularly close contact with the protein. Altogether, the protein makes polar contacts with more than two-thirds of the phosphates in the DNA-binding region. The interdomain β -region holds the DNA in a vice-like grip, penetrating the deepened major groove to varying extents, and making intimate contacts with several bases therein. So this β -region is responsible for recognition of the Ter sequence, as well as for tight binding by Tus. Notably, the tight-binding region is asymmetrically positioned towards the left extremity of the complex, as shown in Fig. 2.

A replication fork moving from right to left is arrested upon encountering the B (blocking) end of the complex. The extent to which Tus grasps and surrounds the DNA enables its rightmost face to present a physical barrier to the helicase. This means that the helicase cannot disrupt the tight Tus-Ter contacts. When moving from left to right, towards the P (passage) end, the helicase encounters no barrier, and proceeds to strip off the relatively exposed red strand and disrupt the Tus primary binding site. Furthermore, the unwinding motion of the helicase could relax the hold of Tus on the DNA by levering back its flanking domains and thereby causing its displacement. This model is given more credence by the fact that 13 Tus mutants which lowered or abolished fork arrest *in vivo* showed reduced, or no, Ter binding. Most of the mutations were mapped to the interdomain region

involved in DNA binding.

So Kamada *et al.*³ see Tus as being essentially a polar clamp on the DNA. A possible molecular basis for such a clamp has been described⁹ in terms of a stable barrier to unwinding at the blocking end because of more extensive crosslinking between Tus and the two DNA strands at that end. But the Kamada model is crucially different — the blocking end of Tus can prevent the helicase from reaching the tight-binding region because Tus protrudes beyond this region, and because of its configuration

around the DNA.

The observed structure itself does not rule out the possibility of interactions between Tus and the helicase or some other protein involved in the replication process, and the model won't necessarily satisfy the proponents of the requirement for such an interaction. However, the crystal structure will enable them to test their views critically through detailed probing of the surface features that are required at the blocking end of Tus for fork arrest.

How might these findings apply to the *B. subtilis* replication terminator protein (RTP)-Ter system? Although the Kamada model would certainly strengthen a case against the involvement of specific RTP-helicase interactions in fork arrest, it is possible that the situation in *B. subtilis* is quite different from that in *E. coli*. Tus and RTP have no sequence similarity, and neither do the DNA terminators of the two organisms. Also, the functional *B. subtilis* arrest complex requires an interaction between two RTP dimers bound to a larger terminator¹⁰. But the *E. coli* picture does raise the possibility of a similarly configured steric barrier arising from the apparently more intricate situation in *B. subtilis*. □

R. G. Wake is in the Department of Biochemistry, University of Sydney, New South Wales 2006, Australia.

- Hill, T. M. in *Escherichia coli and Salmonella* 2nd edn (eds Neidhardt, F. C. *et al.*) 1602-1614 (ASM, Washington DC, 1996).
- Hill, T. M. *Annu. Rev. Microbiol.* **42**, 603-633 (1992).
- Kamada, K., Horiuchi, T., Ohsumi, K., Shimamoto, N. & Morikawa, K. *Nature* **383**, 598-603 (1996).
- Lee, E. H., Kornberg, A., Hidaka, M., Kobayashi, T. & Horiuchi, T. *Proc. Natl Acad. Sci. USA* **86**, 9104-9108 (1989).
- Khatiri, G. S., MacAllister, T., Sista, P. R. & Bastia, D. *Cell* **59**, 667-674 (1989).
- Hiasa, H. & Marians, K. J. *J. Biol. Chem.* **267**, 11379-11385 (1992).
- Sista, P. R., Hutchison, C. A. & Bastia, D. *Genes Dev.* **5**, 74-82 (1991).
- Coskun-Ari, F. F., Skokotas, A., Moe, G. R. & Hill, T. M. *J. Biol. Chem.* **269**, 4027-4034 (1994).
- Gottlieb, P. A. *et al.* *J. Biol. Chem.* **267**, 7434-7443 (1992).
- Langley, D. B., Smith, M. T., Lewis, P. & Wake, R. G. *Mol. Microbiol.* **10**, 771-779 (1993).

Deeply deluded

WE see things in three dimensions because our two eyes have a useful horizontal separation. Each eye sees nearby objects in a different position against the background of more distant ones; the brain compares the two images and derives a view with depth.

Random-dot and repeat-unit stereograms, those popular illusions in many books and posters, subtly exploit this mechanism. They have a pattern of many slightly different units repeated across the page, designed to subvert the eyes' visual fusion of the image. The right eye may focus on unit *n* while the left eye focuses on the very similar unit *n* + 1. This misfusion then propagates across the whole stereogram. The units are shaped and positioned so that this false correlation gives a consistent impression of a scene in depth. Prolonged staring is often needed to see the effect.

The illusion would be more immediate if the eyes could be prevented from making their initial, correct matching between the two views. So Daedalus is doing it. He points out that a diffraction grating appears a different colour from different directions (look at a credit-card hologram). He is devising a computer-controlled holographic laser-printing engine to generate stereograms with diffractive repeat units. Left and right eyes will view each unit from a different angle, and see it in a different colour. The left and right images will be taken to arise from different objects, and will fail to fuse. Instead, fusion will occur between adjacent units, designed to look the same colour to left and right eyes. Prolonged staring will not be needed: the depth illusion will spring to the eye, convincingly and at once.

To defray the expense of the research, the first applications will be military. DREADCO's stereographic camouflage will make tanks and guns look much farther away than they really are, or give them the depth outlines of trees or buildings. Commercial products will soon follow. Stereographic clothes will augment (or diminish) desired regions of the wearer's figure; stereographic wallpaper will give the tiniest flat the illusion of spacious grandeur.

But nature may have got there first. Many animals have vertical stripes or other units which repeat horizontally, but not quite exactly. Such striking patterns, says Daedalus, cannot conceal the creature; they must camouflage it stereographically. He is now wandering round the local zoo, gazing abstractedly at zebras and tigers and angelfish, waiting for them to seem suddenly more distant, or to dissolve into a pattern of unconnected shapes.

David Jones

Paul Erdős (1913–96)

"THE Mozart of mathematics" — who would be worthy of such an epithet? Paul Erdős, who died of a heart attack on 20 September at a conference in Warsaw, was given this appellation, and justifiably so. He was one of the most dazzling minds of our century: when he was barely 20, the great Issai Schur of Berlin dubbed him the "magician of Budapest"; others called him the "occidental Ramanujan" and the "modern-day Euler".

An altogether remarkable man, Erdős lived for his subject with a passion and devotion rare even among scientists. His publications, about 1,500 papers with close to 500 co-authors, outstrip by almost an order of magnitude those of even his most prolific contemporaries. Moreover, it was not just in sheer volume that he excelled. He liked to remind his friends that in the old Hungarian parliament, the votes were not counted but *weighed*; in the same way, a mathematician is judged on the quality of his best theorems. In more than six decades of astonishing activity, Erdős made fundamental contributions to number theory, probability theory, real and complex analysis, geometry, approximation theory, set theory and, especially, combinatorics. Perhaps his genius shone brightest in number theory and combinatorics. He practically created the areas of probabilistic number theory, partition calculus for infinite cardinals, extremal graph theory and the theory of random graphs.

Erdős was born on 26 March 1913 in Budapest, into a Jewish Hungarian family. Both his parents taught mathematics and, until he entered the university of Budapest in 1930, he was educated mostly at home. As a second-year undergraduate he proved his first results, which brought him international acclaim and an invitation to join the remarkable team of mathematicians in Manchester led by Louis Mordell. He had intended to go to Germany, but the rise of fascism prevented him. Jokingly, he suggested that the traditional Jewish toast should be changed to "Next year in Göttingen!". For four very productive years, Erdős stayed in Manchester, with frequent visits to Cambridge, Oxford and London.

In 1938–39 he had a remarkable year in Princeton, at the Institute for Advanced Study: with Mark Kac and Aurel Wintner he founded probabilistic number theory; with Paul Turán he proved important results in approximation theory; and he also gave an unexpected solution to the then outstanding problem in dimension theory. Inexplicably,

his fellowship at the institute was not renewed, and he started travelling. He did not return to Europe until 1948, when he was able to visit Hungary for the first time in more than a decade, to meet his mother and the few members of his family who had survived the Holocaust.

In 1949, Atle Selberg and Erdős gave an elementary proof of the prime number theorem, first proved in 1896 with



the aid of complex function theory. For more than 50 years it had been assumed that an elementary proof could not be given. The 1950s saw great results with Arye Dvoretzky and Shizuo Kakutani on random walks and brownian motion. In the 1960s, Erdős worked much in set theory, the "paradise created by Georg Cantor"; with Richard Rado and András Hajnal he founded partition calculus, a sophisticated study of the relative sizes of large infinite sets. In 1966, with John Selfridge, he solved a century-old problem in number theory. His collaboration with Alfréd Rényi launched the theory of random graphs.

He never had a permanent post, and never applied for one. Even when he had a research position for a year, which was rare, he would frequently visit friends elsewhere to do joint work. "Another roof, another proof" was his legendary motto, and from his twenties he hardly ever slept in the same bed for seven consecutive nights. From the 1960s, he lived mostly in the United States, visiting Budapest most summers, but he also spent periods in Israel, Canada and England.

Erdős has influenced mathematics in several important ways. He proved repeatedly that elementary methods have their place in mathematics. Here 'elementary' does not mean 'simple': an elementary proof relies on ingenuity (which frequently leads to formidable complications), rather than on high-powered, sophisticated tools.

He was the first to fully realize and

exploit the power of random methods to attack a great variety of problems that had nothing to do with random choices. By now it is a commonplace in mathematics that suitable random choices may well produce objects that are painfully difficult to construct explicitly, and the efficiency of many modern practical computer algorithms rests on their use of such 'randomization'. That such a technique is now widespread is due to Erdős.

But it was as a *source* of problems that Erdős was in a class of his own: in the entire history of mathematics there is nobody remotely like him. He has left behind hundreds of attractive problems that are easy to state but that usually turn out to have pinpointed the heart of the matter. From the 1950s, he offered monetary rewards for solutions to his problems, reflecting his estimate of their difficulty. But solving an 'Erdős problem' has always brought more glory than financial gain.

Throughout his life, he was eager to help mathematicians, especially young ones (the prodigy Lajos Pósa being his favourite). There are dozens of mathematicians all over the world who practically owe their careers to him. I was not yet 15 when he introduced me to combinatorics, and so set the course of my life.

"Property is a nuisance", Erdős claimed, and lived very modestly all his life. He did not covet academic honours, and remained somewhat outside the mathematical establishment. Nevertheless, in 1984, he shared the Wolf prize with Shiing-shen Chern, and then gave away most of the \$50,000 he had received. In 1973 the London Mathematical Society elected him an honorary member and in 1975 he was a visiting fellow of Trinity College, Cambridge. He was a member of the national academies of Hungary (1956), the United States (1979), India (1988), England (1989) and others, and received many honorary degrees. It is fortunate that in 1993 George Csicsery made a documentary film about him.

Paul Erdős lived as he wanted to: he "proved and conjectured" to the great benefit of mathematics. He kept up the flood of exciting results well into his seventies, and remained phenomenally productive until the day he died.

Béla Bollobás

Béla Bollobás holds positions at Trinity College, Cambridge, and the University of Memphis, and is currently in the School of Mathematics, Institute for Advanced Study, Olden Lane, Princeton, New Jersey 08540, USA.

Sperm selection by females

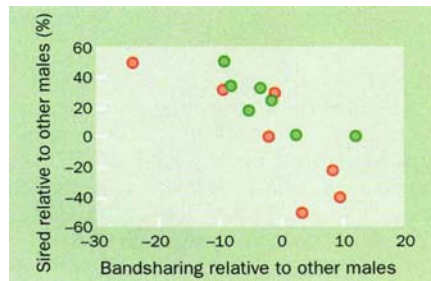
SIR — The history of sexual-selection studies reveals an increasing recognition of the active role of females in determining microevolutionary trajectories. Early studies emphasized male tactics such as combat and scramble competition, and doubted the evolutionary significance of female choice¹. Even after female choice was convincingly documented, its role was believed to be restricted to pre-copulatory phenomena^{1–3}. Despite an increased recent focus on sperm competition within a female's reproductive tract, the female has typically been viewed as providing the arena of competition, rather than being an active participant in the selective process⁴. Our studies on lizards provide the first clear evidence of active selection of sperm by females, in ways that enhance female fitness.

In the Swedish sand lizards (*Lacerta agilis*) we have studied, brother–sister matings reduce offspring viability (and hence maternal fitness)⁵. However, female sand lizards accept copulations from virtually every male that courts them; thus, they often mate with close relatives^{5,6}. Most females mate with more than one male, so natural selection should favour females (or eggs) with the ability to discriminate among the sperm of these competing males, and preferentially use the sperm of distantly related males for fertilization. Observations that multiple copulations with different partners enhance the viability of a female's offspring are consistent with such a mechanism, but are controversial^{7–10}, and have not demonstrated selective use of sperm by females. To do this, we need to know paternity for each offspring.

Our behavioural and DNA fingerprinting studies of a natural population of sand lizards confirm that more closely related males sire fewer offspring per copulation than do more distantly related males. At our study population on the west coast of Sweden, these small lizards (up to 90 mm body length, 20 g) mate in spring; females produce a single clutch of 4 to 15 eggs each year⁵. We observed mating directly in the field, and used the proportion of shared DNA fingerprint bands as an index of genetic similarity⁶.

We found that males with a higher genetic similarity to their partner sire a lower proportion of her offspring than do more distantly related males. This effect is most easily seen if the data are standardized, such that each clutch has the same mean value for relatedness of males to the female, and paternity is calculated in relative terms (the observed proportion of offspring sired by a given male minus the proportion expected from chance, based on the number of males involved: $r = -0.85$, $n = 7$, $P < 0.02$; see figure). This

result is corroborated by staged matings between pedigree-unrelated lizards in the laboratory. Again, we find a strong negative correlation between a male's genetic similarity to his partner (proportion band-sharing) and the proportion of her offspring that he sires ($r = -0.84$, $n = 7$, $P < 0.02$; see figure). Analysis of covariance shows that the relationship between a male's genetic similarity to his mate versus his proportional paternity of the



The proportion of offspring fathered by a male sand lizard, within a multiply sired clutch, depends on the degree of his genetic similarity with the female parent. Males that are genetically more similar to their partners (relative to other males) father a lower proportion of offspring, both in natural matings in the field (green circles) and in controlled matings in the laboratory (red circles). The proportion of offspring sired is calculated as the observed proportion minus that expected if paternity was random (50% if two males contributed to the clutch; 33% if three males contributed). The index of bandsharing is calculated as the male's genetic similarity with the female, minus the mean genetic similarity of the female with all males mating with her. We use information from only one male per clutch, to avoid pseudoreplication.

resulting clutch is similar in matings conducted in the laboratory versus the field (slopes $F_{1,9} = 0.30$, $P = 0.60$; intercepts $F_{1,10} = 4.92$, $P = 0.06$; see figure), with a strong overall effect of the degree of bandsharing on the proportional paternity of the clutch ($F_{1,10} = 23.10$, $P < 0.0008$).

This result might also arise if males that produce more sperm fertilize higher numbers of offspring per copulation^{4,11}. However, proportional paternity is not correlated with either the male's body size (and hence testis size), nor the duration he has for sperm replenishment since his last copulation^{4,11} ($P > 0.40$ for both variables for both laboratory and field data), and remains correlated with male–female relatedness in a partial correlation analysis in which both of these variables are held constant ($r = -0.64$, $n = 16$, $P < 0.03$). Hence, we conclude that female sand lizards actively select the sperm from distantly related males. Even if they are unable to avoid mating with close rela-

tives, females 'choose' not to use sperm from these matings.

This kind of cryptic female choice via intrauterine sperm competition has important implications. Estimates of inbreeding frequency based on behavioural data (numbers of copulations) may be in significant error, for example. More generally, female sand lizards can exert significant mate choice irrespective of their copulatory behaviour. If such abilities are widespread, they may substantially modify our understanding of the determinants of variance in reproductive success in natural populations, and the degree to which female animals can control male reproductive success. Our data suggest that female organisms are capable of subtle discrimination among potential mates, even after copulation, and hence may control microevolutionary phenomena to a greater degree than has heretofore been suspected.

Mats Olsson, Richard Shine

Thomas Madsen

*School of Biological Sciences A08,
University of Sydney,*

New South Wales 2006, Australia

Annica Gullberg, Håkan Tegelström

Department of Genetics,

University of Uppsala,

Box 7003, 750 07 Uppsala, Sweden

1. Andersson, M. *Sexual Selection* (Princeton Univ. Press, 1994).
2. Partridge, L. & Halliday, T. in *Behavioural Ecology, an Evolutionary Approach* (eds Krebs, J. R. & Davies, N. B.) 222–250 (Blackwell, Oxford, 1984).
3. Parker, G. A. *Biol. Rev.* **45**, 525–567 (1970).
4. Birkhead, T. & Møller, A. P. *Trends Evol. Ecol.* **8**, 100–104 (1993).
5. Olsson, M., Gullberg, A. & Tegelström, H. *J. Evol. Biol.* **9**, 229–242 (1996).
6. Olsson, M., Gullberg, A. & Tegelström, H. *Biol. J. Linn. Soc.* (in the press).
7. Madsen, T., Shine, R., Loman, J. & Håkansson, T. *Nature* **355**, 440–441 (1992).
8. Olsson, M., Gullberg, A., Tegelström, H., Madsen, T. & Shine, R. *Nature* **369**, 528 (1994).
9. Olsson, M., Madsen, T., Shine, R., Gullberg, A. & Tegelström, H. *Nature* **372**, 230 (1994).
10. Keller, L. *Nature* **372**, 230 (1994).
11. Simmons, L. W., Stockley, P., Jackson, R. L. & Parker, G. A. *Behav. Ecol. Sociobiol.* **38**, 199–206 (1996).

CO₂ fluctuation at high latitudes

SIR — Keeling *et al.*¹ recently showed that an increased seasonal amplitude of atmospheric CO₂ concentration in northern latitudes is correlated with increasing land temperature, with this correlation being strongest at high latitudes. Because atmospheric CO₂ has been drawn down by plant photosynthesis earlier in the spring in recent years, and because most high-latitude warming has occurred in winter and spring, rather than in summer², Keeling *et al.*¹ suggest that a temperature-driven increased length of growing season accounts for the increased seasonal amplitude of atmospheric CO₂. Here we provide evidence from field measurements that

supports and extends their hypothesis.

In Alaskan arctic tundra, experimental warming causes several plant species to begin growth earlier in the growing season³, with an associated increase in production and biomass of vascular plants⁴. These results are consistent with the hypothesis¹ that earlier snowmelt could increase plant production in early spring and, therefore, shift the timing of seasonal fluctuation in atmospheric CO₂.

Increasing seasonal amplitude of atmospheric CO₂ in response to winter/spring warming could also reflect greater microbial (heterotrophic) respiration in winter, as suggested by observations in northern Russia. Half of the ecosystem respiration in forest tundra of northeast Siberia occurs in winter, probably from microbial respiration in deep soils⁵. If extrapolated to boreal needle-leaved forest, this winter respiration accounts for 40% of the seasonal amplitude in atmospheric CO₂ north of 60° N (ref. 5). These measurements were made in 1989–91, the time of highest Northern Hemisphere land temperatures during the past 30 years¹. Similar high rates of winter respiration have been measured near Barrow, Alaska, the main site of high-latitude atmospheric CO₂ monitoring⁶.

Finally, net primary production is the difference between photosynthesis, which occurs primarily in summer, and plant respiration, which occurs throughout the year with lower rates in winter. Increased plant respiration during winter, which was not considered by Keeling *et al.*¹, could increase the amplitude of atmospheric CO₂ without any change in summer net primary production. Thus, there are several mechanisms by which warmer winters² could increase the seasonal amplitude of atmospheric CO₂ at high latitudes.

F. S. Chapin III

Department of Integrative Biology,
University of California,
Berkeley, California 94720, USA

S. A. Zimov

North-East Scientific Station,
Pacific Institute for Geography, Far-East
Branch,

Russian Academy of Sciences,
Republic of Sakha,
Yakutia, Cherskii, Russia

G. R. Shaver

Ecosystem Center,
Marine Biological Laboratory,
Woods Hole, Massachusetts 02543, USA

S. E. Hobbie

Department of Biological Sciences,
Stanford University,
Stanford, California 94305, USA

Gene acquisition in HIV and SIV

SIR — Recombination is a fundamental feature of retroviruses¹ and is important in the evolution of human and simian immunodeficiency viruses². Recombinant examples of HIV-1, HIV-2 and SIV_{AGM} have been found², including some with widespread geographical dissemination^{3,4} (subscripts defined in Fig. 1 legend). In all these cases, the recombination event involved exchange of homologous genomic regions. Here we describe the first example of a non-homologous recombination event which resulted in the acquisition of a new gene by SIV_{SM}, the sooty mangabey virus which gave rise to HIV-2.

The known primate lentiviruses form five main groups² (Fig. 1a). Members of the SIV_{SM}/HIV-2 lineage have two genes (*vpx* and *vpr*) within the *vif-tat* genomic region, whereas all other primate lentiviruses contain just one (usually termed *vpr*). It has been suggested that *vpr* and *vpx* share sequence similarity, and thus a common ancestry^{5,6}. Taking the latest consensus *Vpr* and *Vpx* sequences⁷ from diverse HIV and SIV, we have pro-

duced an alignment which confirms this similarity (Fig. 2). Furthermore, our recent finding that the *vpr* and *vpx* genes of SIV_{SM} each have separate functions which are both encoded by the *vpr* gene of HIV-1 supports their relatedness⁸. The simplest (and currently accepted) model to explain the presence of the additional gene in SIV_{SM}/HIV-2 is a gene duplication event after this lineage diverged from the others^{5,6} (see Fig. 1a). If this were true, *vpr* and *vpx* from SIV_{SM}/HIV-2 would be expected to be more closely related to each other than to *vpr* from any other lineage. However, we have performed phylogenetic analyses (Fig. 1b) which contradict this: in fact, *vpr* from SIV_{SM}/HIV-2 is very similar to *vpr* from HIV-1/SIV_{CPZ}, whereas *vpx* from SIV_{SM}/HIV-2 is most similar to the *vpr* of SIV_{AGM} from African green monkeys.

Although the *vpr/vpx* tree (Fig. 1b) must be treated with some caution, the clustering of SIV_{SM}/HIV-2 *vpx* with SIV_{AGM} *vpr* (rather than with SIV_{SM}/HIV-2 *vpr*) is fairly robust. Thus, although these sequences are difficult to align at their termini, we have focused the analysis on two stable regions of the alignment (Fig. 2). These regions are short, but if the highly divergent sequence from SIV_{SYK}

is excluded a high bootstrap value is found for the branch separating the sequences of HIV-1/SIV_{CPZ}, SIV_{SM}/HIV-2 (*vpr*) and SIV_{MND} from those of SIV_{AGM} and SIV_{SM}/HIV-2 (*vpx*). This is clearly not consistent with a duplication event after the divergence of the SIV_{SM}/HIV-2 lineage. Moreover, while there is no outgroup available to root the *vpr/vpx* tree definitively, it is clear that no alternative rooting would cluster the *vpr* and *vpx* sequences of SIV_{SM}/HIV-2 together to the exclusion of sequences from other viruses.

An alternative explanation — a gene duplication before the initial radiation of the primate lentiviruses — can be also excluded for several reasons: it would require multiple independent gene deletion events across the various lineages; any divergence of function between *vpr* and *vpx* that occurred before deletion of one of them would have to be compensated by the remaining gene reacquiring that function; and the apparently close relationship between SIV_{SM}/HIV-2

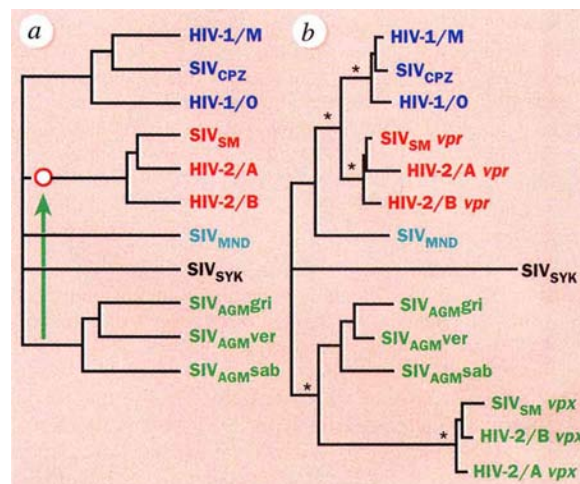


FIG. 1a, Phylogenetic relationships among primate lentiviruses, derived from Pol protein sequences (adapted from refs 2, 3): there are five main lineages, whose branching order is unresolved. The diversity of HIV-1 is indicated by examples of groups M and O; diversity of HIV-2 by subtypes A and B. SIV subscripts denote species of origin: CPZ, chimpanzee; SM, sooty mangabey; MND, mandrill; SYK, Sykes' monkey; ver, vervet; gri, grivet; sab, sabaeus; the last three are subspecies of African green monkey (AGM). The circle indicates the approximate position of the *vpr/vpx* gene duplication event under the old hypothesis^{5,6}; the arrow indicates the *vpx* gene acquisition event proposed here. b, Phylogenetic relationships among *vpr* and *vpx* sequences. The tree has been rooted so as to be most congruent with the relationships among these viruses indicated by their Pol sequences (part a). The tree was derived by neighbour-joining analysis of protein sequence (PAM) distances implemented using PHYLIP¹⁰; results using minimum evolution, maximum parsimony, and maximum likelihood methods did not differ significantly. Branches found in at least 80% of bootstrap replicates using all four methods are indicated by an asterisk (SIV_{SYK} was excluded from the bootstrap analyses).

- Keeling, C. D., Chin, J. F. S. & Whorf, T. P. *Nature* **382**, 146–149 (1996).
- Chapman, W. L. & Walsh, J. E. *Bull. Am. Meteorol. Soc.* **74**, 33–47 (1993).
- Chapin, F. S. & Shaver, G. R. *Ecology* (in the press).
- Chapin, F. S. *et al. Ecology* **76**, 694–711 (1995).
- Zimov, S. A. *et al. Clim. Change* **33**, 111–120 (1996).
- Oechel, W. C., Vourlitis, G. & Hastings, S. J. *Global Biogeochem. Cycles* (in the press).

HIV-1 Vpr	MEQAP---EDQGPQREPYNEWLELEELK?EAVRHFPRIWLHSLGQHIYET---
HIV-2 Vpr	MTEAPTE?PP---EDE?PQREPWDEWV?EVLEEIKEEALKHFDPRLL?ALGNYYI?R---
HIV-2 Vpx	MTDPREVPVPGNSGEETIGEA?F?--WLDRTVEEINR?AVNHLPRELIFQVWQSW?YVHD
SIV _{AGM} Vpr	MASGRDPRE??PGWLEIWDL?REPWDEWLRDML?DLNQEA?HFGNRLFRVWNYCVVEG?R

HIV-1 Vpr	-YG-DTWEGVEAIRILQQLFIHFRIGCQHSR-----IGITRQRARRNGS--SRS
HIV-2 Vpr	-HG-DTLEGARLIRILQRLAF?HFRAGC?HSR-----IGQGGGNPLSAIPPSRGMO
HIV-2 Vpx	EQGMS?SYTKYRYLCLMQKAMF?H?KKGCRCLG-----GGHGGGWRPGPPPPPPGLA
SIV _{AGM} Vpr	-?G?P??E??YKYRIRIVQKALFVHFRGCGRRR?PF?PYEERR?GQGGG?A??P?PGL
	* * * * *

FIG. 2 Alignment of consensus Vpr and Vpx sequences of HIV-1, HIV-2 and SIV_{AGM}. Bold asterisk, residues conserved across all four sequences; asterisk, residues conserved between HIV-2 Vpx and (at least) one other sequence. Residues highlighted in bold in the HIV-2 Vpx sequence are sites where the three other sequences are identical, and may represent adaptive changes (see text). The overlined regions were used for the phylogenetic analyses (Fig. 1b).

vpx and SIV_{AGM} is not explained.

In fact, the tree (Fig. 1b) strongly suggests that the vpx gene in SIV_{SM}/HIV-2 did indeed appear after the divergence from other primate lentiviruses, but by acquisition of the SIV_{AGM} vpr gene (Fig. 1a) rather than by duplication of SIV_{SM}/HIV-2 vpr. This would have required recombination between the ancestral SIV_{SM} and SIV_{AGM} lineages, which could only have occurred if a single monkey was coinfecting with both viruses. African green monkeys (of the sabaues subspecies) and sooty mangabeys share habitats in West Africa, and we have previously reported evidence of recombination between their viruses³, involving the replacement of the 3' gag and 5' pol genes in SIV_{AGM}sab with the corresponding regions from an ancestral SIV_{SM}. The event proposed here is quite distinct, in that we are invoking the addition of a gene to an ancestral SIV_{SM} by non-homologous recombination. Taken at face value, the tree would suggest that the donor was an ancestral SIV_{AGM}. However, the position of SIV_{AGM}sab on the same branch as SIV_{AGM}ver and SIV_{AGM}gri is not strongly supported; in some analyses SIV_{AGM}sab was found to switch positions to the SIV_{SM}/HIV-2 lineage. Thus, the source of the vpx gene could have been SIV_{AGM}sab.

Changes in the HIV/SIV genome structure (rearrangement or addition of genes) are usually detrimental to viral replication, yet it is implicit in the above that the recombinant SIV_{SM} must have replaced the existing virus in sooty mangabeys, suggesting that a selective

advantage may have been conferred by the additional gene. Whereas HIV-1 needs both Vpr and the gag-encoded matrix nuclear localization domain to establish a productive infection in terminally differentiated macrophages⁹, SIV_{SM} appears to require only Vpx⁸. Thus, through the acquisition of the SIV_{AGM} vpr gene, SIV_{SM} may have been able to replicate more efficiently in non-dividing target cells (for example, Langerhans cells in the mucosa) and achieve enhanced transmissibility. (There could be other advantages to having two proteins with specialized functions.) Branch lengths in

the tree suggest that vpx evolved very rapidly after acquisition. This was probably not due to relaxation of selective constraints, since the rates of evolution of vpx and vpr are now similar (judged by the diversity within SIV_{SM}/HIV-2), but may reflect an initial period of adaptation of vpx to a new role in SIV_{SM}.

Paul M. Sharp*

Elizabeth Bailes

Department of Genetics,
University of Nottingham,
Queens Medical Centre,
Nottingham NG7 2UH, UK
e-mail: paul@evol.nott.ac.uk

Mario Stevenson

Program in Molecular Medicine,
University of Massachusetts Medical
Center,

Worcester, Massachusetts 01605, USA

Michael Emerman

Division of Molecular Medicine,
Fred Hutchinson Cancer Research Center,
Seattle,
Washington 98104, USA

Beatrice H. Hahn

Departments of Medicine and Microbiology,
University of Alabama at Birmingham,
Alabama 35294, USA

*To whom correspondence should be addressed.

Turbulence and financial markets

SIR — Ghashghaie *et al.* discuss analogies between the price dynamics in the foreign exchange market and three-dimensional fully developed turbulence¹. We have independently carried out a study comparing the dynamical properties of the S&P 500 index and of the time evolution of a three-dimensional fully turbulent fluid, but we draw different conclusions. Specifically, we find that although intermittency — abrupt changes of activity in the time evolution of the variance of price changes and of the mean energy dissipation — and non-gaussian behaviour (for short times) in the probability distribution of price and velocity changes characterize both systems, the stochastic nature of the two processes is quantitatively different.

Among the differences we find are: first, price changes of the S&P 500 are substantially uncorrelated (we detect only a slightly enhanced diffusion), but velocity changes in three-dimensional turbulence are anticorrelated. Specifically, we find that the measured autocorrelation function of the price changes is a fast-decaying monotonic function with a characteristic time of a few minutes and the spectral density of the index is a power-law $S(f) \propto f^{-2}$ for more than four orders of magnitude, but in three-dimensional turbulence an inertial range $S(f) \propto f^{-5/3}$ exists (Fig. 1).

Second, the maximum of the probability distribution $P(Z = 0)$ of price changes $Z_{\Delta t}(t)$ after a time Δt shows clear Lévy

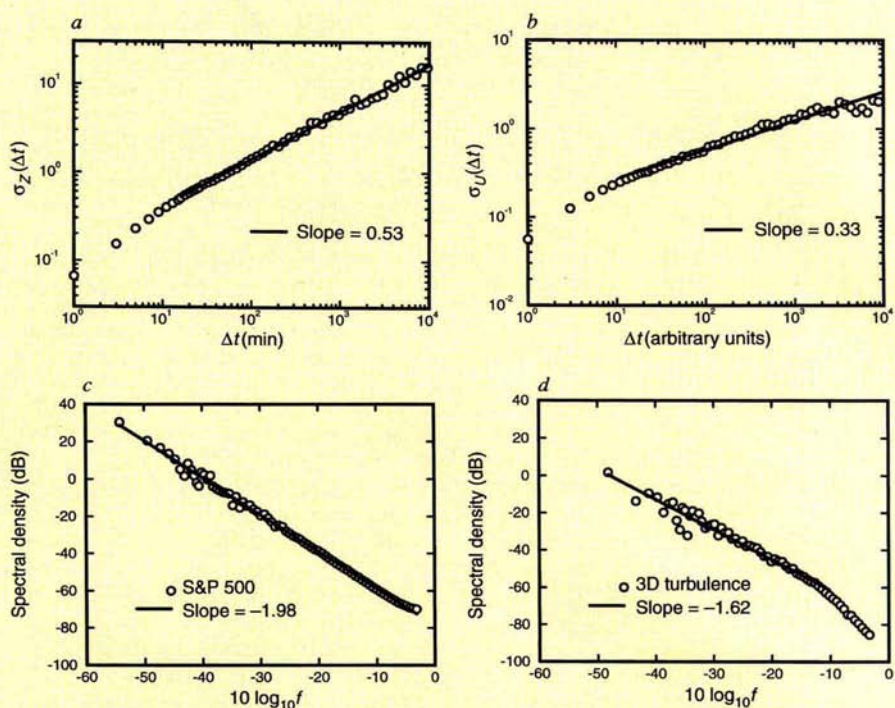
(non-gaussian) scaling (for $1 \leq \Delta t \leq 1,000$ minutes) as a function of Δt (ref. 2), but the corresponding turbulence quantity does not show scaling (Fig. 2).

One interesting result presented in ref. 1 is that the distribution of price changes in the foreign exchange market is changing shape as a function of the time delay Δt . The distribution evolves from a leptokurtic (with tails fatter than in a gaussian distribution) to a gaussian shape. Although a similar behaviour is observed in fully developed turbulence, the deep difference observed in the correlation (of price and velocity changes) and scaling (of the central part of the respective distributions) properties of the two processes leads us to conclude that indeed, at the moment, the simplest model describing the main features of the dynamics of prices in speculative markets is the truncated Lévy flight (TLF) introduced in refs 2–4.

In this model, independent identically distributed price changes are characterized by a Lévy stable distribution in the central part of the distribution, but a cutoff is present after which the distribution is not power-law. This truncation implies that the process is characterized by a finite variance. A TLF is only roughly self-similar and the process eventually converges to a gaussian distributed process due to the finiteness of the variance, as predicted by the central limit theorem. The TLF also describes quite well the dynamics

- Coffin, J. M. in *Fields Virology* (eds Fields, B. N. & Knipe, D. M.) 1437–1500 (Raven, New York, 1990).
- Sharp, P. M., Robertson, D. L. & Hahn, B. H. *Phil. Trans. R. Soc. Lond. B* **349**, 41–47 (1995).
- Jin, M. J. *et al.* *EMBO J.* **13**, 2935–2947 (1994).
- Gao, F. *et al.* *J. Virol.* **70**, 7013–7029 (1996).
- Tristem, M., Marshall, C., Karpas, A., Petrik, J. & Hill, F. *Nature* **347**, 341–342 (1990).
- Tristem, M., Marshall, C., Karpas, A. & Hill, F. *EMBO J.* **11**, 3405–3412 (1992).
- Myers, G. *et al.* (eds) *Human Retroviruses and AIDS 1995* (Los Alamos National Laboratory, New Mexico, 1996).
- Fletcher, T. M. *et al.* *EMBO J.* (in the press).
- Heinzinger, N. K. *et al.* *Proc. Natl Acad. Sci. USA* **91**, 7311–7315 (1994).
- Felsenstein, J. *PHYLIP* (Phylogeny Inference Package) version 3.5c (Dept of Genetics, Univ. Washington, Seattle, 1993).

FIG. 1a, Standard deviation $\sigma_Z(\Delta t)$ of the probability distribution $P(Z)$ characterizing the price changes $Z_{\Delta t}(t)$ plotted double logarithmically as a function of Δt for the S&P 500 time series. After a time interval of superdiffusive behaviour ($0 < \Delta t \leq 15$ minutes), diffusive behaviour close to the one expected for a random process with independent identically distributed increments occurs; the measured diffusion exponent 0.53 is very close to the theoretical (uncorrelated) value $1/2$. b, Standard deviation $\sigma_U(\Delta t)$ of the probability distribution $P(U)$ characterizing the velocity changes $U_{\Delta t}(t)$ plotted double logarithmically as a function of Δt for the velocity-difference time series in turbulence. Data recorded in the atmosphere at a Taylor microscale Reynolds number R_λ of the order of 1,500 (data provided by K. R. Sreenivasan). After a time interval of superdiffusive behaviour ($0 < \Delta t \leq 10$), a diffusive behaviour close to the one expected for a fluid in the inertial range is observed (the measured diffusion exponent 0.33 is close to the theoretical (anti-correlated) value $1/3$). c, Spectral density of the S&P 500 time series for the time period 1984–87 representative of the 6-year time period 1984–89 (an investigation performed for the time period 1986–89 gives a curve overlapping with the figure shown). The $1/f^2$ power-law behaviour expected for a random process with independent increments is observed over a frequency interval of more than 4 orders of magnitude. d, Spectral density of the velocity time series of a three-dimensional fully developed turbulent fluid. The $1/f^{5/3}$ inertial range (low frequency) and the dissipative range (high frequency) are clearly observed.



of a price in foreign exchange markets (A. Arneodo *et al.*, preprint cond-mat/9607120 at <http://xxx.lanl.gov>). Features observed in economic data that are not explained in terms of the TLF model are the time dependence of the scale factor parameter γ of the TLF distribution, which shows a fluctuating behaviour with bursts of activity localized in limited intervals of time, and the long-

range correlations observed in the time evolution of γ and in the variance of price changes (two quantities related to what is called 'volatility' in the economics literature) (ref. 5).

Can we reconcile the known intuitive parallels between finance and turbulence with the fact that quantitatively the two phenomena are quite different? This quantitative difference might disappear

for a turbulent system in an abstract space of non-integral dimensionality. Theoretical studies^{6,7} show that for $d \approx 2.05$ there exists (under the Taylor hypothesis) an uncorrelated turbulent behaviour characterized by a spectral density $S(f) \propto f^{-2}$, just as for the S&P 500. Further study is required to test if a 2.05-dimensional turbulence could be really consistent with the stochastic properties of market data and with the main assumption of mathematical finance that is that no arbitrage is possible in an efficient market⁸.

Rosario N. Mantegna

Istituto Nazionale di Fisica della Materia,
Unità di Palermo & Dipartimento di
Energetica ed Applicazioni di Fisica,
Università di Palermo,
Palermo I-90128, Italia

H. Eugene Stanley

Center for Polymer Studies and
Department of Physics,
Boston University,
Boston, Massachusetts 02215, USA

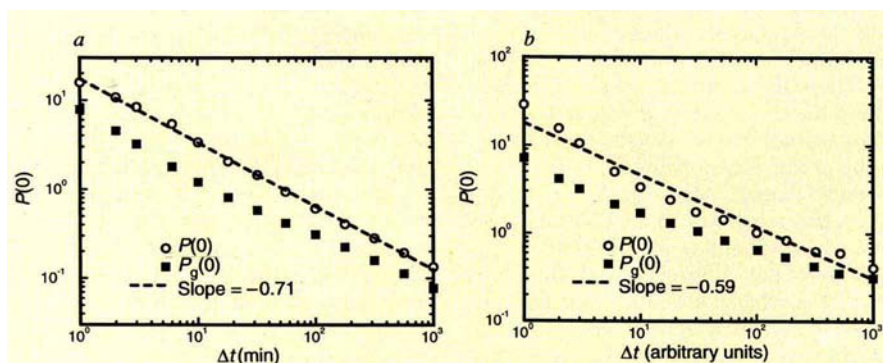


FIG. 2a, Probability of return to the origin $P(Z=0)$ for the S&P 500 index (circles) and $P_g(Z=0) = 1/\sqrt{2\pi}\sigma(\Delta t)$ (filled squares) as functions of the time sampling interval Δt . $P_g(Z=0)$ is the probability of return to the origin expected for a gaussian stochastic process determined by measuring the standard deviation $\sigma(\Delta t)$ of the experimental data. The two measured quantities differ in the full interval implying that the profile of the probability density function (PDF) must be non-gaussian. A power-law behaviour is observed for the entire time interval spanning three orders of magnitude. The slope of the best linear fit is -0.71 ± 0.025 . The difference between the two quantities is decreasing when Δt increases, implying a convergence to a gaussian process for high values of Δt . b, Probability of return to the origin $P(Z=0)$ (circles) and $P_g(Z=0)$ (filled squares) (defined in a) as functions of the time-sampling interval Δt for the velocity of the fully turbulent fluid. Again, the two measured quantities differ in the full interval, implying that the profile of the PDF must be non-gaussian. However, in this case, a single scaling power-law behaviour does not exist for the entire time interval spanning three orders of magnitude. The slope of the best linear fit (which is of quite poor quality) is -0.59 ± 0.11 .

1. Ghoshghale, S. *et al.* *Nature* **381**, 767–769 (1996).
2. Mantegna, R. N. & Stanley, H. E. *Nature* **376**, 46–49 (1995).
3. Mantegna, R. N. & Stanley, H. E. *Phys. Rev. Lett.* **73**, 2946–2949 (1994).
4. Mantegna, R. N. & Stanley, H. E. in *Lévy Flights and Related Topics in Physics* (eds Shlesinger, M. F., Zaslavsky, G. M. & Frisch, U.) 300–312 (Springer, Berlin, 1995).
5. Mantegna, R. N. & Stanley, H. E. *Scaling Approach to Finance* (Cambridge Univ. Press, in the press).
6. Frisch, U., Lesieur, M. & Sulem, M. *Phys. Rev. Lett.* **37**, 895–897 (1994).
7. Fournier, J.-D. & Frisch, U. *Phys. Rev. A* **17**, 747–762 (1978).
8. Samuelson, P. A. *Industrial Management Review* Vol. 6 No. 2, 41–49 (1965); reproduced as Chapter 198 in Samuelson, P. A. *Collected Scientific Papers* Vol. 3 (MIT Press, Cambridge, MA, 1972).

Fruit and nuts

Walter Gratzer

Vitamanía: Vitamins in American Culture. By Rima D. Apple. *Rutgers University Press: 1996. Pp. 245. \$48 (hbk); \$18.95 (pbk).*

In 1911 Casimir Funk enriched the language with a new word: *vitamine* was a corruption of 'vital amine' and therefore appropriately enough a misnomer. It engorged at once the imaginations of all the mountebanks and entrepreneurs who preyed on a boundlessly gullible public, animated, as Rima Apple observes in her social history of the subject, by a potent mix of fear, hope and guilt.

Dr Johnson described the career of a notorious quack of the day (the infamous John "Chevalier" Taylor) as "an example of how far impudence will carry ignorance". A number of the Chevalier's successors feature in Apple's story. In modern times the United States has, in the view of at least one authority on the subject, James Harvey Young, provided the most fertile ground for medical quackery. Among many more malign examples, he has chronicled the success in recent memory of Silent George of Shawneetown, Illinois, who got rich by buying tins of condensed milk, spraying them with gold paint, and selling them, relabelled "Swamp Rabbit Milk", "a balanced product for unbalanced people, rich in Vitamins J U M and P", and not wholly devoid of aphrodisiacal properties. Such were often in fact imputed to vitamins generally.

In proclaiming the virtues of vitamin preparations, the self-deluded were perhaps more dangerous than the charlatans, on the familiar principle that fools are worse than villains, who at least sometimes take a rest. Apple relates, for instance, how Dr Russell N. Wilder of the Mayo Clinic, no less, fed a group of "sociable, contented workers" a diet deficient in thiamine; within a few weeks their personalities were hideously transformed, for they became quarrelsome, depressed and tired, and (horror!) "they even went on strikes". When thiamine was restored to them they all became "their old selves again".

Evidence grounded in such quantitative rigour proved irresistible, and vitamin supplements went on sale in chain stores, such as Woolworths. With enough vitamins, an advertiser insisted, you would have bright eyes, but with an insufficiency, dull eyes; you could choose between lustrous hair or dull or falling hair, good bones or poor bone structure, a "fit feeling" or tiredness and irritability, "buoyant vitality" or the other thing. Nor was this all: Dr Norman Hayhurst Jolliffe of the New York Bellevue Hospital satisfied himself that polyneuritis

in alcoholics was caused not by alcohol but by vitamin B deficiency, for he had made the acquaintance of a barman who had consumed 15 fluid ounces of whisky every day for forty years with no ill-effects. The reason was evident: it had ever been his custom to top up his restorative with milk, which contained vitamin D as an antidote. (This had an echo quite recently when a baseball player who advertised a popular breakfast cereal on television was accused of rejecting the product himself. "Sure I eat wheaties", he responded, "a bowl of wheaties with bourbon for breakfast can't be beat.")

Much of the vitamin advertising was aimed at the softest part of the market, mothers with young children, whose insecurity the manufacturers exploited pitilessly. "Inside, where it can't be seen, the damage starts!" ran the Squibb cod liver oil advertisement, "A defective development of the bone structure so insidious that it is more than likely to touch even the well-cared-for baby in intelligent modern homes!". Terror as an instrument of marketing is no longer generally acceptable (except in politics), but in 1926 its efficacy was not in doubt: one of the most successful promotions was based on a picture of a sinister cluster of masked and gowned surgeons looking down at something unmentionable, over the caption "And it all began with harsh toilet tissue."

Vitamin retail sales in the United States increased from a bare \$700,000 in 1925 to \$82.7 million by 1939, and by 1937 the consumption of cod liver oil had reached 5.79 million gallons, enough to float the US Pacific fleet. Vitamins, Apple suggests, had dragged science out of the laboratory and planted it at the centre of American culture (or should it be folklore?) and business.

Between a credulous American public and the unscrupulous assault of the food and pharmaceutical industries stood a lone sentinel — the Food and Drug Administration (FDA). With so much at stake (vitamin sales at the last count stood at \$4 billion), the lawsuits raged, hinging on such issues as whether vitamins were food or therapeutic agents, as most of the advertising had implied. Similar battles have been fought over such deceptions on the sick as the spurious cancer remedy laetrile: its inventor's case was based on the contention that his nostrum

was a vitamin, and cancer a disease of dietary deficiency.

Just as proof of any proposition, however preposterous, can be found somewhere in the medical literature, so an expert can be procured to support any claim. There is always then of course an antipodal expert, eager to testify to the contrary. The FDA succeeded in moderating the more extravagant assertions of large manufacturers such as Miles Laboratories. They were also able to drive a few shifty smaller organizations out of business, most notably in 1961 the Pannett Company, which marketed the highly successful Acnotabs — vitamin pills guaranteed to eliminate acne.

Dear Old GRANDMA MEANT Well...



*Say, Auntie! I've
done a smart deed, didn't I?
and the whole has been coming
to me as good old Ole Grandpa
meant well.*

*I'm from some other place
nearly every body has a copy
this morning. The year is only
one month old. It's a fine
new thing. But don't you
mind. We'll have it soon.
I'll give you one.*

Grandma

**BUT... She Had
Never Heard About VITAMINS**

- If she had, she never would have written the letter. You see, Grandma accepted the common idea and frequent notion of her day; as something you had in "you and bear". Fewer poor teeth and improperly formed bones in the children were things the "lads was heir to". Nothing could be done about it. So—why worry?
- But... Something was done about it. What men can dream of—Science can do. How well that has been proven by the discovery and success of Vitamins! It took years of course—years of hard work and heartbreak, but those "hungriest" of the laboratory knew that it would come. And it did come! What folly of a generation ago accepted as "fate"—we recognize take in stride—thanks to the advent of Vitamins.
- NOW we know how essential Vitamins are in our everyday life and we also know that two of the important recognized Vitamins are A and D. Vitamin A is necessary

in building normal resistance to infection more likely to occur as membranes of the eyes, ears, nose, throat and skin, so when there is a deficiency in this Vitamin Vitamin D helps in proper growth and bone size defects. We need an essential for normal tissue formation in children.

- But then came the problems of high prices and unavailability. Families rich and poor had as equal right to the benefits of this great discovery, but first people had to go to the new Dr. Miles Laboratories, Inc., developed ONE A DAY (brand) Vitamin A and D Tablets—a seasonal forward step in economy and convenience.
- No longer is it necessary for you or your family to risk the expense caused by Vitamin A and D deficiency. Prices are now—30 Tablets are only .35¢, 30 Tablets only .50¢ and 100 Tablets only \$1.50. ONE A-DAY Tablets are easy to take, pleasant to take and are biologically standardized to insure normal daily requirements.

ONE A DAY IS ALL YOU NEED! ONE PENNY A DAY IS ALL IT COSTS!—So give your family the benefits of purest Vitamin A and D all year long! Start today with ONE A DAY (brand) Vitamin A and D Tablets. Ask your druggist for them by name—ONE A DAY. Look for the big "M" on the package.



ONE A DAY
VITAMIN A and D TABLETS



MILES LABORATORIES INC., ELKHART, IND.

Scientific claims used to sell vitamins in this advertisement from 1941.

In recent years, however, the FDA has come under attack from the political right in venal alliance with the food industry. It has been denounced as "paternalistic", a term these days of abuse rather than approval, and accused of "treating the American people as children". In 1976 its powers to protect the public from dishonest claims were curtailed by an act of Con-

gress for the first time since the food and drug legislation came into force in 1906. Deregulation in the interests of "consumer choice" (that is to say, in this case, the food industry) is a familiar phenomenon of recent years, and has already yielded such dividends as "mad cow disease". The deterioration of diet through poverty and the prevalence of processed foods (amounting to some 70% of all food-stuffs sold) has now made supplementary vitamins a necessity for many, so the food industry may have won its battle in any case.

Apple has put together a scholarly and absorbing tale. Her prose lapses from time to time into *Timespeak* ("backward ran sentences until reeled the mind", as Woolcott Gibbs put it), but the narrative carries the reader along briskly enough. Yet this is a book with a serious purpose and has lessons which, if we do not absorb, we shall certainly be destined to repeat, about the manipulation and abuse of science by industry, by opportunistic mountebanks and by politicians; for the enemy lurks in every cranny of our free-market society. □

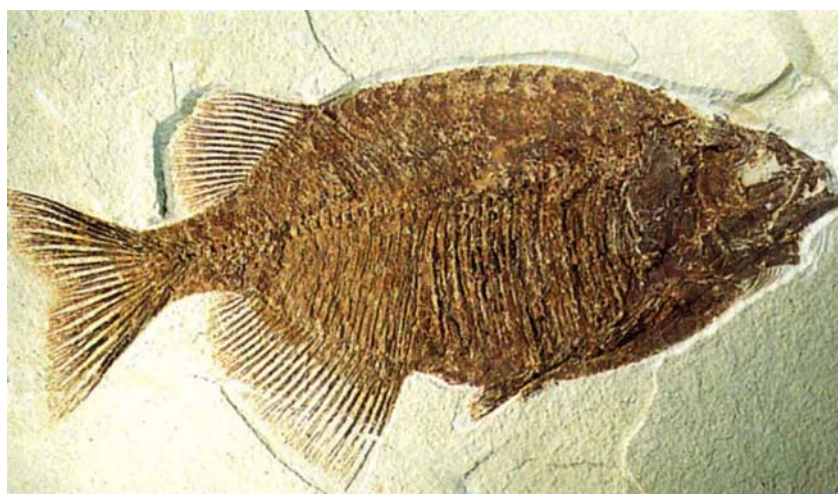
Walter Gratzer is in the MRC Muscle and Cell Motility Unit, King's College, 26-29 Drury Lane, London WC2B 5RL, UK.

Evolution in the past tense

Axel Meyer

Darwin's Dreampond: Drama on Lake Victoria. By Tijs Goldschmidt. (Translated by Sherry Marx-Macdonald.) MIT Press: 1996. Pp. 274. £21.50, \$25.

LAKES Victoria, Malawi and Tanganyika in East Africa are home to adaptive radiations of cichlid fish. These are the world's most diverse assemblages of vertebrates, with each species flock comprising hundreds of endemic cichlids. Lake Victoria is the youngest of the three lakes, with an



A FOSSIL of *Phaerodus testis*, an osteoglossid from the Eocene Green River shales in Wyoming, USA. From *The Rise of Fishes: 500 Million Years of Evolution* by John A. Long, out now in paperback. John Hopkins University Press, \$65 (hbk), \$34.95 (pbk).

estimated age of less than half a million years. It is also the largest tropical lake in the world, having a surface area almost twice the size of the Netherlands. It contained (notice the past tense) a flock of more than 300 species of haplochromine cichlids.

It has recently been suggested that climate change led to its basin drying up completely less than 15,000 years ago. This raises the possibility that evolution does not take "a sea of time", but that most of the mud-biters, algae-scrapers, leaf-choppers, snail-crushers, snail-shellcrackers, zooplankton-eaters, insect-eaters, prawn-eaters, fish-eaters, paedophages (eggs and young cichlids are brooded in the mouth of their mother and are sometimes sucked out by these child-eaters), scale-scrapers and parasite cleaners found there arose astonishingly quickly from a small number of ancestral species. None of the other groups of fish or invertebrates in the lake have radiated to any comparable degree.

How can so many species have arisen within such a short period of time? Since the cichlids were discovered more than a hundred years ago, biologists have puzzled over the mechanisms of their evolutionary origin and ecological maintenance. Charles Darwin did little to elucidate the actual mechanisms of speciation, although he knew about the beautiful colours of cichlids, which suggest that sexual selection must have played a big role in their evolution. If he had travelled to Lake Victoria instead of the Galápagos Islands, would he have said anything different about evolution and speciation?

The demise of Lake Victoria's cichlids was the largest

mass extinction of vertebrate species in human history, but has received little attention, perhaps because cichlids do not have cute blue eyes and brown hair.

Tijs Goldschmidt lived on Mwanza Gulf in Tanzania from 1981 to 1986, when the disaster of extinction unfolded. He took on the sad task of documenting it, in a perversely entertaining and non-judgmental way, in this eloquent book. But describing doom is not what Goldschmidt set out to do. Rather, he tried to tackle the question of how so many species could coexist within one, admittedly extraordinarily large, body of water. This was a formidable task considering that most of the 300 or so species of cichlids have still not been formally described. He left the phylogenetic analysis of cichlids to "gene freaks", "bright-eyed and bushy-tailed molecular biologists", who are currently sequencing genes from cichlids of the East African lakes.

Goldschmidt was part of the Dutch haplochromine ecology survey team that devoted more than 20 years to collecting baseline ecological, morphological, taxonomic and fisheries data on the cichlids, which provide the basic protein source for the people around Lake Victoria. These data were to serve as a biological database for the planned Dutch-aided fisheries for these cichlids.

A couple of decades before the Dutch arrived, however, the non-native, gigantic, top-of-the-food-chain Nile perch was released into the lake in Uganda. This wilful but later forgotten introduction, and the more recent addition of the water hyacinth, had disastrous ecological consequences on the lake's ecosystem. The forests around Lake Victoria were cut down to smoke the meat of the Nile perch, so there was increased erosion. This was one factor in the lake becoming eutrophic, with the result that large portions of it periodically



A dangerous perch for Lake Victoria's cichlids.

became deoxygenated, leading to the deaths of huge numbers of fish.

Goldschmidt noted the ever-increasing changes in species composition, and the eventual demise of more than two-thirds of all endemic cichlid species. Most of the cichlids that ate open-water fish and zooplankton became extinct, a total of about 200 species. This is the best-documented case study of a mass extinction, and is unique both in terms of sheer numbers of species and the speed of their vanishing. Western notions of time rarely apply to Africa, but here Noah's Ark leaked and sank remarkably quickly. Despite several attempts to save the cichlids, such as captive-breeding programmes and plans to create areas of the lake free of Nile perch, this fight is lost. This evolutionary theatre, where one of the most diverse faunas on earth played, is closing, turning foul, and being overgrown with water hyacinth.

Goldschmidt's book is both unusual and eminently readable. He alternates between personal narrative told in the first person (such as the anatomical search for the essence of the laugh in the muscles of the human head) and scientific passages, which are sometimes enigmatically selected. For example, on the same page we can read both the amusing history of how a Nazi SS cap reached the lake and a scientific account of various species concepts in evolutionary biology. Evolutionary principles, such as frequency-dependent selection, are explained in parallel with the human experience.

In Africa, where malaria reigns and sickle-cell anaemia can provide a selective advantage, fieldwork can be difficult; it can also be rich in interactions with people who often have wondrously different philosophies. Goldschmidt's daily research is like the work of Sisyphus and, by his own account, consisted mostly of catching fish and thrusting a thermometer deep into the lake. He suggests that if he did this work often enough, he "would come to understand how evolution worked".

His work was influenced by the politics of President Julius Nyerere, the search for a bridal dowry, and malaria. Unfortunately, there is little time to ponder evolutionary theory while struggling to find two meals per day. The often Kafkaesque nature of African bureaucracy is beautifully exemplified by a tale that describes the task of finding out the dimensions of the narrowest tunnel on the railway from Dar-es-Salaam to Mwanza, in an attempt to bring a Dutch commercial fishing vessel to Lake Victoria. It becomes clear "that everything can and cannot pass through that tunnel".

Just as the cichlids are in imminent danger of extinction, so too are the taxonomists, who are "roaming around the margins of science" and are the only ones who know the names of endangered species. Several members of the Dutch

team were later unable to find employment and, despite Goldschmidt's excellent ecological and taxonomic publications, he has left science and is now pursuing a writing career. □

Axel Meyer is in the Department of Ecology and Evolution, State University of New York, Stony Brook, New York 11794-5245, USA.

An X-pert chemist

W. H. Brock

Edward Frankland: Chemistry, Controversy and Conspiracy in Victorian England. By Colin A. Russell. *Cambridge University Press: 1996. Pp. 535. £65, \$110.*

EDWARD Frankland's greatest contribution to nineteenth-century chemistry was the formulation of the concept of chemical bonding, or valency. As an eminent teacher and examiner for the government's Department of Science and Art, he included the notion of the combining powers of atoms into the emerging science educational system of the 1860s.

A crony of Huxley, Tyndall, Hooker and Spencer, he was also one of the formidable band of Victorian experts and exponents of scientific naturalism who gave scientific evidence and advice in legal and parliamentary inquiries. As a member of the X-Club, he conspired to exert influence on the science policies of the Royal Society and the Chemical Society, as well as on the latter's professional spin-off, the (Royal) Institute of Chemistry, which he helped to create in 1877.

Although now largely forgotten, Frankland's most significant contribution to public life was probably the cleansing of the nation's water supplies, through his development in 1867 of a powerful method of water analysis and his work as a virtually solo commissioner on river pollution. Before the introduction of bacteriological assessment of water, it was Frankland who ensured that domestic water supplies were clean and safe by challenging private water companies with what he called, controversially, the "previous sewage contamination" of water.

He was knighted in 1897, but has remained until now a little-known and little-studied Victorian scientist. This is despite the fact that his contribution to the transformation of Victorian science and society was as important as that of Edwin Chadwick, John Simon or Charles Darwin, and that his theoretical contribution to chemistry was of fundamental help to August Kekulé in the creation and dissemination of structural theory.

Colin Russell first drew attention to Frankland's construction of a theory and

language of valency, bonding and graphic formulae thirty years ago in his definitive *History of Valency* (1967). Russell was then a professional chemist engaged in research on heterocyclic chemistry, but interest in Frankland's career transformed Russell into a historian of science, and launched him on his career at the Open University.

In 1986, after a decade studying archives in Lancashire in northwest England, during which an extensive collection of Frankland's papers and letters still in private hands were discovered, Russell published *Lancastrian Chemist: The Early Years of Sir Edward Frankland*. This engagingly written mystery story told in rich local detail of Frankland's illegitimate birth to a Lancashire servant girl in 1825, his disjointed education, his apprenticeship to a Lancaster pharmacist in 1840, and how he was encouraged to learn chemistry. Destined for a medical career, he qualified in neither pharmacy nor medicine, and in 1845 the local medical grapevine found him employment as an analyst in Lyon Playfair's laboratory at the Museum of Economic Geology in London.

Lancastrian Chemist ended at this point with the promise of a second volume that would deal with Frankland's education in Germany and his later career, first in Manchester from 1851 to 1857, and finally in



Taken from *Edward Frankland*

Frankland: formulated valency theory

London, as a pre-eminent government chemist. Sensibly, however, *Edward Frankland*, this latest, very readable and well-illustrated book, is a complete biography, and incorporates the essential features of the earlier book while not replacing it entirely for serious research purposes. *Lancastrian Chemist* showed that no single factor in Frankland's fascinating Lancastrian upbringing — even the fact that his natural father, Edward Gorst, paid for his education — could explain his later success. The detailed exploration of that academic and financial success is the heart of this new biography.

In Russell's view, Frankland's drive and ambition were the result of his psychologi-

cal need for social legitimacy. However, a major factor in his success was undoubtedly the rigorous training in gas analysis that he received from Robert Bunsen and his pupil Hermann Kolbe at the University of Marburg, where in 1849 Frankland was the first Englishman to take a PhD. Frankland quickly came to admire German culture and its notion of class based on mind rather than wealth. He was to have it both ways.

Russell provides a fascinating account of chemistry teaching at Owens College, which Frankland ultimately found lacking in challenge, and of his wide-ranging work as a chemical consultant. He had married Sophie Fick, whom he met at Marburg, and his commitments to a growing family drove him into industrial consultancy. Although this gave him financial security, there is evidence that he often compromised himself; he consequently left Manchester with his reputation as an expert witness somewhat clouded. Despite his academic renown, Frankland's permanent connection with 'trade' demeaned him in the eyes of scientific contemporaries, including the other members of the X-club, who denied him the presidency of the Royal Society.

Frankland moved to London in 1857 and, after holding a variety of positions, he finally found scope for his commitment to the communication of science at the Royal College of Chemistry, in succession to A.W. Hofmann. By this time, Frankland had made his name as a great synthetic organic chemist. Russell does not duck some difficult organometallic and synthetic chemistry, but confines it to two technical chapters, and by deft summary in other contexts he provides an overall account that can be understood and enjoyed by scientists from other disciplines, as well as by historians of science.

In 1866, Frankland published *Lecture Notes for Chemical Students*; like the examinations he set for the Department of Science and Art, this used the concept of valency he had developed in the 1850s and the graphic use of chemical formulae. Graphic formulae are now such an indispensable part of chemical language for making clear the constitution and structure of compounds that it is difficult to imagine the revolutionary nature and significance of Frankland's innovation.

Frankland became a member of the X-Club in the 1860s, and came to regard some fellow members, Huxley (and by extension Darwin) and Tyndall, as the modern evangelists who would guide men's thoughts long after the four gospel writers had become obsolete. However, Frankland's religious views — youthful evangelical conversion followed by doubts — were more complicated than simple agnosticism or atheism.

Frankland's family relationships were similarly tortuous. His first wife died in

1873; two years later he married Ellen Grenside, who was not much older than his two daughters. His children were openly hostile to the match. Because all four of Sophie's children kept diaries that have survived, Russell is able to reconstruct this painful episode in detail. And Frankland's relationship with his son Percy (who also became an eminent chemist) deteriorated so far that father and son were not on speaking terms for several years.

Such vivid details of Frankland's private life and fascinating insights into his relationships with other eminent Victorians make this a rich and rewarding read. Frankland was undoubtedly one of Britain's most important and influential nineteenth-century scientists, and it is good to see him properly honoured at last with a fine biography. □

W. H. Brock is in the Department of History, University of Leicester, Leicester LE1 7EH, UK.

Mind of a machine

William Clocksin

Impossible Minds: My Neurons, My Consciousness. By Igor Aleksander. *Imperial College Press: 1996. Pp. 347. £17.*

MANY of the goals of research into machine intelligence have been achieved by biological nervous systems, but using radically different strategies, architectures and hardware. For example, in living neural systems there are no formal or numerical representations; communication media are stochastic; events are asynchronous; components are unreliable and widely distributed; connectivity does not obey precise blueprints; and 'clocking' speeds are millions of times slower than in digital computers. Yet the performance of natural systems in real-time tasks entailing perception, learning and motor control, in unpredictable and perilous environments, remains unrivalled.

How can they be so effective? Can any underlying principles be exposed and incorporated within new artificial systems? Can artificial systems designed on the principles of electronics approach the performance of natural systems?

Igor Aleksander seeks to answer these questions, and his book provides a popular account of the problem of consciousness. It is a personal view, from the perspective of a plain-speaking and clear-sighted engineer, articulated with all the optimism and confidence that engineers need to pursue the audacious goal of designing intelligent and conscious robots.

The core of the book is constructed around three interleaved stories. The first

is about Magnus, a suite of software being developed by Aleksander and his colleagues to establish whether state machines with a feedback path from output to input can display complex behaviours and 'mental' operations. Magnus also serves as the engineer's ever-present focus on the artefact. The hypothetical Magnus of the future is described and represented with the same conviction as the primitive work-in-progress Magnus that runs on a desktop computer in London. The story culminates with a hypothetical conversation between Magnus and a journalist in the year 2030.

A parallel science-fiction story is presented with one episode at the end of each chapter. It describes a world in which robotic life-forms debate whether humans are conscious.

But the main story is built around Aleksander's "basic guess", which supposes that consciousness is determined by the firing patterns of neurons. This is followed by thirteen consequences, dealing with topics such as representation, language learning, instinct and emotions. The technical issues unfolded in this story are given in terms of state-machine automata with feedback and a novel iconic representation of percepts and memories. That might sound daunting, but readers do not need to know details of the theory or implementation.

This book is different from many popular books on consciousness in that there is emphasis on the social life of the robot as a useful input for conditioning its experience and learning language. The robot's principal contact with memories and behaviours other than its own comes from what Aleksander calls the "societal repository". Although this attention to culture and society represents an advance on current views in the machine-intelligence community, it is clear to me that a more radical approach is called for, in which the mind and consciousness are a construction of the organism's engagement with its cultural and social matrix. So instead of Kelly's personal construct theory, the main influence from psychology in this book, perhaps designers of the next Magnus should turn for psychological inspiration to the Milan school of systemic practice, itself inspired, ironically enough, by cybernetic principles.

This book touches on philosophy, psychology, cybernetics, artificial intelligence and electrical engineering, but does not engage with any particular issue in much technical depth. Rather, it is more of an 'ideas' book, and was fun to read. It should be an adventure and an inspiration for the human, if not yet the robotic, mind. □

William Clocksin is at the Computer Laboratory, University of Cambridge, Cambridge CB2 3QG, UK.

Molecular indicators of secondary oil migration distances

S. R. Larter*, B. F. J. Bowler*, M. Li*[†], M. Chen*, D. Brincat*, B. Bennett*, K. Noke*, P. Donohoe*, D. Simmons[‡], M. Kohnen[‡], J. Allan^{§||}, N. Telnaes[¶] & I. Horstad[#]

* Fossil Fuels & Environmental Geochemistry (Postgraduate Institute), Newcastle Research Group (NRG), Drummond Building, University of Newcastle, Newcastle upon Tyne NE1 7RU, UK

[‡] KSEPL, Shell, Rijswijk, The Netherlands

^{||} Imperial Oil Resources, Calgary, Canada

[¶] Norsk Hydro, Bergen, Norway

[#] Saga Petroleum, Sandvika, Norway

The ability to determine the distance migrated by an oil from source rock to reservoir could greatly assist in the identification of new, economically viable accumulations of petroleum. Non-alkylated benzocarbazoles, which are present in trace quantities in oils, exhibit changes in both absolute and relative concentrations that correlate with migration distance, independent of the maturity of the oils. Such molecules appear promising as empirical—but quantitative—indicators of the relative migration distances of oils, and represent a first step towards true source–reservoir distance indicators.

PETROLEUM systems consist of a source rock, where petroleum is generated through the thermal alteration (maturation) of buried organic matter, a carrier bed, and a trap with the reservoir and seal where the petroleum accumulates¹. Secondary petroleum migration, in which typically oil or gas moves buoyantly up and away from the site of generation in the source rock, through an inclined carrier bed to the trap, occurs on lateral scales of up to several hundred kilometres (ref. 2) and remains the least understood yet the most critical of the processes involved in petroleum accumulation. In the search for new petroleum resources, information is required about the route and distance travelled by migrating oils away from their source rocks but this information cannot be explicitly provided by seismic or other remote-sensing surveys. Thus, molecular indicators of the absolute or relative distance migrated by oils away from their source rock have been sought for decades³.

For economic reasons, direct sampling of petroleum systems is usually only possible through boreholes located in the traps, so assessments of source-rock and carrier-system properties from the geochemical analysis of petroleum in reservoirs are of great value. The determination of source-rock organic-matter type and maturity (a parameter related to the maximum temperature experienced by a source rock at the time of oil expulsion) through the analysis of fluid samples in reservoirs has been widely applied^{4,5}. Some success in defining regional petroleum migration pathways from the application of biomarker hydrocarbon geochemistry has been achieved^{1,6} but the successful application of geochemistry to the definition of secondary migration distance has not, despite bold efforts⁷.

The use of tracer molecules, determined by the analysis of reservoired oils and associated gases, to quantify petroleum migration processes has been attempted using a variety of strategies. Ideally, tracers should initially exist in only one component of a migration pathway (that is, in the oil, water or solid phases), they should be conserved within the system and they should distribute

predictably into, or from, the oil to the other phases during migration. Determinations of ²⁰Ne and ³⁶Ar partitioning from petroleum system waters into migrating oils have been used to assess relative volumes of interacting oil and water in a North Sea oil accumulation⁷, and changes in the concentrations and distributions of alkylphenols, benzene or aromatic nitrogen compounds partitioning from migrating oil to carrier bed waters and rock surfaces have been proposed to monitor oil migration⁸. Yamamoto⁹ and Li *et al.*¹⁰ have suggested that alkylquinolines and alkylcarbazoles present in related oils and source-rock extracts showed varying isomeric distributions and homologue abundances indicative of migration-related fractionations, but they could not distinguish between the effects of primary oil migration out of the source rock and secondary migration in the carrier bed. During migration, alkylcarbazoles undergo systematic fractionations between the various isomers and homologues present in petroleum¹¹, these fractionations being similar to normal phase chromatographic fractionation processes¹². Although the principal fractionations occur when petroleum is expelled from the source rock, less intense fractionations of petroleum alkylcarbazoles also occur during secondary migration, suggesting that aromatic nitrogen compounds may have a potential role as secondary migration tracers.

Although the maturities of different oils reservoired along migration pathways are often quite similar they sometimes vary considerably, reflecting variations in the average temperature of the source rock when the petroleum was expelled. Ideally, migration tracers based on the relative proportions of different isomeric or homologous nitrogen (or other) compounds in oils would not reflect variations in oil maturity but would solely reflect the effect of secondary migration processes. From concerns over the influence of source maturity on the distributions of alkylcarbazoles in migrated oils we were led to consider the non-alkylated benzocarbazoles as possible migration tracers. Whereas maturation-related changes in isomer and homologue distributions of methyl-, dimethyl- or trimethyl-carbazoles in oils occur, the isomeric benzocarbazoles (Fig. 1) are not alkylated and are unlikely to isomerize at oil-generation temperatures; they are therefore less likely to show maturity-dependent behaviour.

[†] Present address: Geological Survey of Canada, Calgary, Canada.

[§] Deceased.

With increasing secondary migration distance, the non-alkylated benzocarbazoles in suites of source-related oils exhibit both a decrease in the ratio of benzo[a]carbazole to benzo[c]carbazole and a reduction in concentration of the two isomers. We propose that the abundance ratio of benzo[a]carbazole/benzo[a+c]carbazole in oils may be used as an empirical, largely maturity-independent measure of the relative secondary migration distance of related oils. We suggest that preferential removal, during migration, of the more rod-shaped benzo[a]carbazole relative to the sub-spherical benzo[c]carbazole is related to selective sorption of benzocarbazoles from the oil onto clay minerals and into solid organic matter in the carrier bed.

Benzocarbazoles in migrated oils

Figure 2 shows the ratio benzo[a]carbazole/(benzo[a]carbazole + benzo[c]carbazole) (the benzocarbazoles ratio (BC ratio)) for data sets derived from the analysis of reservoir oils in five different petroleum systems, two in Western Canada and three in the North Sea. These petroleum systems have been extensively studied, and secondary migration pathways and distances are considered to be relatively well known compared to those in many other basins. The secondary migration distances given in Fig. 2 are our best estimates of the lateral distances travelled by the petroleum beyond a reference oil nearest the source rock. These estimates may have substantial associated errors, especially for the long-distance migrated oils in the Canadian petroleum systems where the detailed paths taken by the oils are not known. The outlying data point at ~290 km migration distance in the Colorado Group sourced oils (Fig. 2), for example, could in fact have migrated much less than 50 km. By comparing non-degraded reservoir oils in different reservoirs along migration routes, variations in chemical composition may be attributed to secondary migration processes and/or to source maturity effects. The influence of primary migration processes is effectively removed from consideration by the choice of samples.

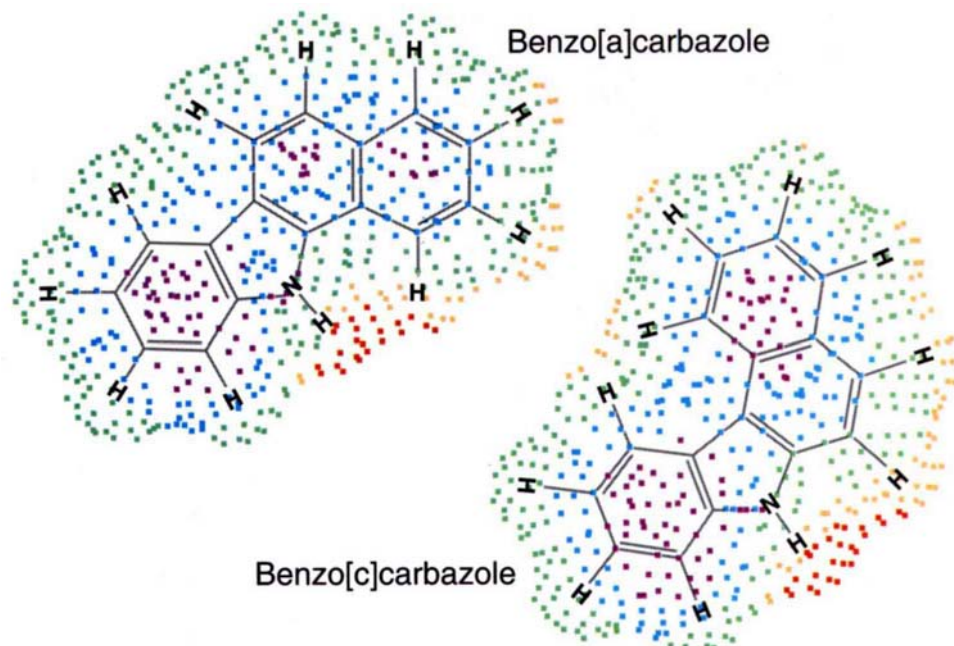
The general trend observed with the combined data sets (Fig. 2) is for the BC ratio to decrease with increasing migration distance. Data sets 1 (Duvernay source) and 2 (Colorado Group source) represent petroleum from two separate petroleum systems in the foreland basin setting of the Alberta basin¹³ in reservoirs at distances of up to 500 km from the mature source rocks. The

long migration distances are typical of foreland basin settings. The BC ratios have large analytical errors for samples beyond ~100 km migration distance where absolute concentrations of benzo[a+c]carbazoles are very low (for example, <0.16 µg per g oil for the carbonate carrier system (set 1; Duvernay) and <0.02 µg per g oil for the clastic carrier suite (set 2; Colorado group)).

Similar broad regional trends, in which the BC ratio decreases with increasing migration distance, are also seen in oil suites from China, South America, the Middle East and Europe. Data sets 3, 4 and 5 (Fig. 2) comprise BC ratio values from North Sea systems characterized by relatively short-distance migration. The simple ratio of benzo[a]carbazole/benzo[c]carbazole ([a]/[c] ratio) also decreases with increasing secondary migration distance and has a predictably larger range of variation, but it is less linear in behaviour (not shown in Fig. 2).

Data set 3 (Central Graben) represents data from a suite of related oils migrating from Upper Jurassic Kimmeridge Clay Formation source rocks in the Central Graben into Upper Jurassic Fulmar Formation sandstones in the Kittiwake field area of the North Sea, UK sector^{14,15}. Biological marker and aromatic hydrocarbon data indicate that the maturity of the petroleum in this suite increases systematically with decreasing secondary migration distance, a not uncommon occurrence in oils along migration pathways¹⁵. Set 4 (Oseberg South) shows BC ratio data from a suite of related oils derived from Upper Jurassic source rocks which have migrated through Brent group sandstones in the Oseberg South area of the North Sea. The oils show little variation in maturity and are from exploration tests on a possible migration route into one of the large structures in this area. In contrast to these four data sets, data set 5 represents oils of a single petroleum population reservoir at various distances along a possible migration route in the Tampen Spur area of the North Sea. In this case, we initially considered that these oils, sourced in the Upper Jurassic, migrated through Middle Jurassic and Upper Jurassic sandstones over distances up to 25 km (ref. 6). Recent remapping of the area suggests that this interpretation may require revision (I.H., unpublished results), the low absolute concentrations of benzocarbazoles found in these oils implying that they may have migrated over much longer absolute distances than indicated. The absence of a clear trend of decreasing BC ratio with (presumed) increasing migration distance for these oils suggests that we must either reconsider the geological interpretation,

FIG. 1 Connolly dot surface plots²⁸ of benzo[a]carbazole and benzo[c]carbazole. The molecules are oriented with the effective dipole horizontal, the molecular electrostatic potential and partial charges ranging from negative values associated with parts of the nitrogen-bearing ring system (purple and blue) to positive partial charges and molecular electrostatic potentials associated with the pyrrolic hydrogen atom adjacent to the nitrogen atom in the molecule (red).



or be warned that migration does not always produce changes in the relative proportions of benzocarbazoles.

Concentration variations during migration

Concentrations of benzo[a]carbazole plus benzo[c]carbazole in the oils we have examined to date range from 0.02 to over 4 $\mu\text{g/perm g}$ oil and usually decrease with increasing secondary migration distance. In cases where benzo[b]carbazole has been detected, it is usually present in very low concentration and is not included in our 'total' benzocarbazole data. Figure 3 shows the variation in normalized concentration of total benzocarbazoles as a function of relative migration distance for several oil sets, each originating from marine source rocks rich in organic matter of algal/bacterial origin (type II kerogen).

Although the number of case histories shown here is limited, the data suggest that the decrease in concentrations of benzo[a]carbazole plus benzo[c]carbazole occurs predominantly during the first 100 km of petroleum migration, this being a typical maximum migration distance for most basins other than foreland basins. The rate at which benzocarbazoles are removed from oils during migration will, no doubt, be critically dependent on the rate of migration and the nature and composition of the carrier bed itself (for example, fracture versus capillary flow, clastic versus carbonate content, and the contents of clay and organic matter).

Accompanying reductions in the BC ratio, the preferential removal of benzo[a]carbazole relative to benzo[c]carbazole with increasing migration distance is inferred from concentration data for oils on a migration pathway in the Central Graben area of the North Sea (Fig. 4). (The data shown is data set 3, described above.) The oils in this suite show a decrease in maturity with increasing migration distance, with later-expelled, higher-maturity oils being more common in the reservoirs nearest the maturing source rock¹⁵. Maturity variations in the oils delivered to a carrier bed through time might well be expected to complicate the use of benzocarbazole tracers to assess migration distance owing to the expulsion of oils with different initial benzocarbazole compositions into the carrier system. However, our observations to date suggest that such maturation effects are not significant.

Effects of source maturity

As mentioned above, an ideal molecular migration-distance monitor should be influenced solely by migration processes and not, for example, by maturity, which is known to have a marked effect on molecular isomerization reactions. It would be surprising if non-alkylated benzo[a]carbazole and benzo[c]carbazole were related through free isomerization of one to the other at the temperatures involved in oil generation and expulsion in source rocks (100–150 °C) as several bond cleavages and rearrangements would be involved. However, selective generation or destruction of one or other of the isomers in the source rock as it matures could result in variation in the BC ratio of the oil expelled into the carrier. It appears, however, that this is not a serious problem.

A comparison of the BC ratio and oil source maturity (from estimates of source vitrinite reflectance, $\%R_o$, derived from a suite of molecular markers and gasoline range hydrocarbons analysed in the oils) showed no strong relationship for a suite of ten related North Sea oils with source maturities in the range 0.64–1.11 $\%R_o$ (correlation coefficient –0.43; data not shown). This suggests that any variations in BC ratio for oils related to variations in source maturity are small compared to the variation caused by migration. Similarly, no strong relationship is seen between the BC ratio and source maturity for a suite of ten Monterey Formation source rocks with a maturity range of 0.33–0.86 $\%R_o$ (correlation coefficient –0.39; data not shown), while BC ratio data for extracts from a homogeneous Chinese lacustrine oil-prone source rock¹⁶ with a depth range of 2–4 km (immature to oil-window mature) showed a scatter with no clear maturity-related trend. A recently acquired data set from a very homogeneous series of Duvernay Formation source rocks from Alberta, covering the maturity range 0.4–1.5 $\%R_o$, suggests that in this case the BC ratio in the petroleum

within this source rock increases from around 0.5 at 0.5 $\%R_o$ to around 0.7 at 0.8 $\%R_o$, and then shows a scatter from 0.6 to 0.7 over the range 0.8–1.5 $\%R_o$ (ref. 17). Despite wide variations in benzocarbazole concentrations and distributions within source sections, the expelled oils probably have a relatively narrow range of initial BC ratio values. Oils from the Miller Field¹⁸ (North Sea basin) have a similar maturity and source character to mature Upper Jurassic source rocks in the field, and secondary migration distances from equivalent shales off-structure are very small¹⁹. Although a wide range of BC ratios (0.43–0.58) are found in the source rock extracts, the average BC ratio values for these extracts (mean 0.5; $N = 24$) and for the oils (mean 0.54; $N = 9$) are very similar²¹.

Within the maturity range during which most oil is expelled from typical oil-generating source rocks (0.8–1.1 $\%R_o$) BC ratios do not show any consistent source maturity-related trend, but are

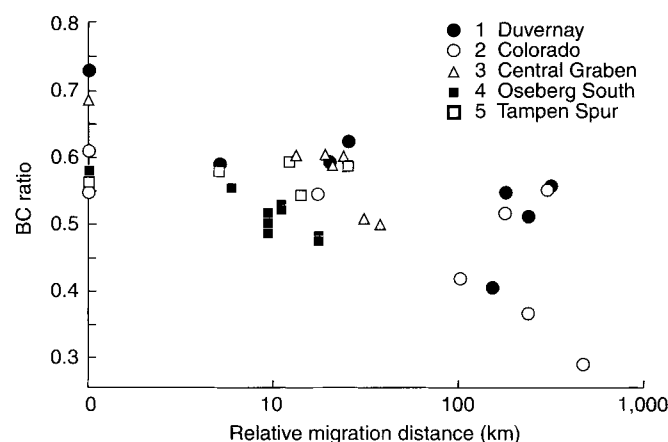


FIG. 2 The ratio of benzo[a]carbazole/(benzo[a]carbazole + benzo[c]carbazole), the BC ratio, for reservoir oils from five petroleum systems plotted versus best estimates of secondary migration distance relative to a reference oil nearest the source rock. The reference oil is given an arbitrary migration distance of 1 km. The oils comprising data sets 1 and 2 are from the foreland basin of Alberta and British Columbia in western Canada. Set 1 represents oils of similar maturity (the $\alpha\beta\beta/(\alpha\alpha\alpha + \alpha\beta\beta)C_{29}$ sterane maturity parameter ($\% \beta\beta C_{29}$ steranes) varying from 0.62 to 0.64 and the hopane $Ts/(Ts + Tm)$ (Ts , 18 α (H)-trisnorhopane; Tm , 17 α (H)-trisnorhopane) parameter varying from 0.63 to 0.75) sourced from marine shales of the Frasnian basal Duvernay facies of the Ireton Formation which are found in dolomitized reefs and banks of the Givetian–Frasnian Devonian Beaverhill Lake–Woodbend–Winterburn groups¹³. Migration distances are variable, and can be short (<1 km) and vertical due to juxtaposition of source rock and trap. Data set 2 represents petroleum of similar maturity ($\% \beta\beta C_{29}$ steranes varying from 0.54 to 0.66; $Ts/(Ts + Tm)$ parameter varying from 0.59 to 0.74) sourced from the Cenomanian–Santonian age marine shales of the Colorado group and mostly reservoir in shallow sands of the Cardium and Viking formations. This petroleum system has a maximum lateral extent of ~500 km between peak-maturity source rock and the most distal oil (Plato)¹³. One oil is slightly biodegraded but we suggest (M.L. and S.R.L.; unpublished data) mild degradation does not affect the ratio. The two petroleum systems are separated by a regional seal^{13,29}. Data set 3 is from a suite of related oils, described in the text¹⁴. Biological marker, gasoline-range and aromatic-hydrocarbon data indicate that the maturity of the petroleum increases systematically with decreasing secondary migration distance. While the C_{30} diahopane/ C_{30} diahopane + C_{29} normoretane maturity parameter varies from 0.86 to 0.71 and the hopane $Ts/(Ts + Tm)$ parameter varies from 0.85 to 0.60 with increasing migration distance, the 4-methyldibenzothiophene/1-methyldibenzothiophene ratio varies from 6.72 to 1.39. Set 4 represents related oils, described in the text. The $\% \beta\beta C_{29}$ sterane maturity parameter varies from 0.62 to 0.72 and the $Ts/(Ts + Tm)$ parameter varies from 0.52 to 0.72. The samples comprising data set 5 represent oils of similar maturity ($\% \beta\beta C_{29}$ steranes in the range 0.52 to 0.57; C_{30} diahopane/ C_{30} diahopane + $C_{29}\beta\alpha$ hopane) in the range 0.47 to 0.60), described in the text. The petroleum system is described by Horstad *et al.*⁶.

generally scattered around initial values of 0.5–0.75 depending on the specific source-rock section studied (corresponding to initial values of 1–3 in the benzo[a]carbazole/benzo[c]carbazole ratio). We suspect that during primary migration, petroleum with an 'average' initial BC ratio for a source basin is probably expelled, despite local variation within the source rock.

Interaction with carrier bed phases

The generally observed decrease in the absolute concentrations of both benzocarbazole isomers in oils migrated over distances of up to 100 km (Figs 3 and 4) suggests that the variation in the benzo[a]carbazole to benzo[c]carbazole ratios during migration results from differential removal of the benzocarbazoles from the migrating oil charge due to interactions with carrier bed phases, especially solid organic phases and minerals (most probably clay minerals). Estimation of the solubility parameter of benzocarbazole using a group type contribution method²⁰ gives a value of 24.9 MPa^{1/2} (in the range for coals), suggesting strong affinity for benzocarbazoles by dispersed organic matter in carrier beds and consistent with previous reports of strong absorption of polycyclic aromatic compounds by coals²¹. Although both benzocarbazole isomers are polar species with similar estimated dipole moments, the rod-shaped benzo[a]carbazole has an effective dipole moment of 1.41 debye units oriented along the long axis of the molecule whereas the sub-spherical benzo[c]carbazole (dipole moment 1.75 debye units) has the dipole oriented closer to the major axis of the molecule (Fig. 1). These differences suggest that solid-phase sorption characteristics and organic solvent solubilities of the two isomers would be quite distinct. Benzo[c]carbazole, for example, has a solubility in a 2:1 mixture of *n*-hexane/toluene of 5.4 mg ml⁻¹ at STP, whereas that of benzo[a]carbazole is 1.3 mg ml⁻¹.

In a laboratory experiment in which oils were flowed through water-wet sandpacks containing added, initially dry, kerogen and clay layers, a reduction in the BC ratio of the oil was observed¹⁹, thereby identifying solid organic matter and clay minerals as candidate sinks for the removal of benzocarbazoles from oils. We have measured sorption characteristics for benzocarbazoles from hexane/toluene (2:1) solution (5 ml) agitated in contact with crushed, dry, brown coal or very illite-rich shale (500 mg), at initial solution concentration levels similar to those found in petroleum. It was clear that no equilibrium situation or endpoint was reached, the removal of benzocarbazoles appearing to continue to beyond the 14-day duration of the experiments. Extrapolation of the sorption data suggests that several months may be necessary for sorption to cease in this system, and whereas both illitic shale and brown coal showed significant uptake of benzocarbazoles from the hydrocarbon phase, the illite-rich material showed a very dramatic selectivity for the removal of benzo[a]carbazole relative to benzo[c]carbazole, even when it had been equilibrated with water vapour. Attempts to desorb compounds from dry illitic shale using fresh solvent, after six months of equilibration with hexane/toluene solutions containing a variety of alkylcarbazoles and benzocarbazoles, resulted in significant recovery of most isomers but negligible recovery of benzo[a]carbazole.

Although the details of the rates and mechanisms of benzocarbazole removal during migration remain unclear, the evidence available to date suggests the involvement of interactions between migrating oils and carrier-bed solid phases rather than chemical destruction of the compound. The uptake of polar compounds by clays is well known²² and it is clear that we must consider kinetically controlled (or even irreversible) processes in the removal of benzocarbazoles from oils during migration. Relatively long breakthrough curves obtained in studies of sorbing tracers moving through porous media suggest kinetically controlled sorption processes²³. If the changes in benzocarbazole concentrations and distributions are due to sorption processes, the 100-km distance scale of reduction in concentrations of benzocarbazoles in migrated oils (Fig. 3) implies kinetically controlled sorption processes which are slow relative to the timescale of petroleum

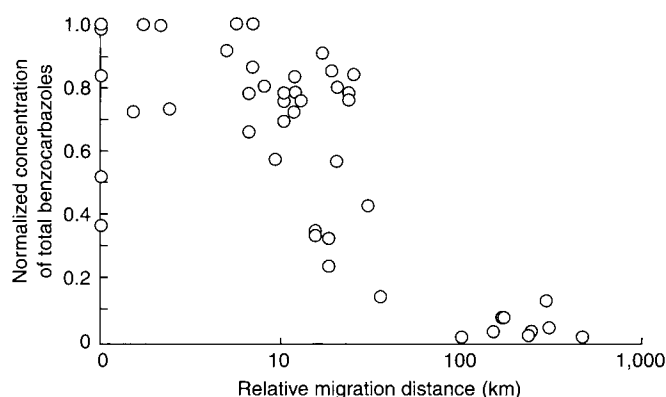


FIG. 3 Variation in the relative concentration of benzocarbazoles as a function of relative migration distance for oils in data sets 1–5 (see Fig. 2 legend) and oils from the Miller Field—all originating from marine source rocks rich in type II organic matter. The Miller field data was generated using the procedure described for data sets 1, 2 and 3. The migration distances for the Miller Field oils are measured from the main fill point of the structure. To facilitate comparison, the concentrations for each data set have been normalized to that of the oil in each data set with the highest concentration of benzo[a]carbazole.

migration. This may imply that diffusion of benzocarbazoles through surface films or zones on sorbing solid phases is the rate-limiting process, with the narrower, more rapidly diffusing benzo[a]carbazole reaching sorbing sites faster than benzo[c]carbazole. Laboratory experiments produce sorption rates which are far too rapid to describe the observed field data used with simple sorption models, indicating that much further study is needed in this area.

Implications of a viable migration tracer

Subject to confirmation by further field studies, the ability to monitor reliably relative oil migration distance by the analysis of reservoir oils would find many applications in both fundamental and applied petroleum geoscience studies.

Here we have addressed lateral oil migration in relatively high-permeability lithologies; less is known about vertical petroleum migration through low-permeability lithologies such as mudstones/shales, yet this process must occur over distances of up to several kilometers in many basins. The rates and mechanisms of vertical fluid migration through mudstone/shale sequences represent primary controls on basin evolution, leakage of oil through shale seals being the most common mechanism of destruction of petroleum accumulations²⁴. However, the processes involved in transmitting petroleum through mudstones/shales are disputed, and the relative importance of fractures or mudrock pore systems as the conduits for petroleum movement is undetermined²⁵. A quantitative mechanistic model of a migration tracer such as the benzocarbazoles may be able to test some of the proposed hypotheses on oil migration through fine-grained rocks by analysing oils migrated to various distances through mudstone/shale sequences. Oils migrated through clay-mineral/organic-matter-rich mudstone pore systems would be expected to show more greatly reduced BC ratios and benzocarbazole concentrations compared to oils having migrated through faults and fractures.

The definition of regional and local migration pathways in many major oil provinces of the world is, in truth, poorly known. Much greater confidence in quantifying the relative distances migrated by oils within part of a basin could aid in the discovery of new oil accumulations by enabling petroleum geologists to assess more accurately the migration pathways used to fill known accumulations. It has been our recent experience that petroleum geologists are often prepared to revise their regional migration models on the basis of benzocarbazole data which indicate, for example, that

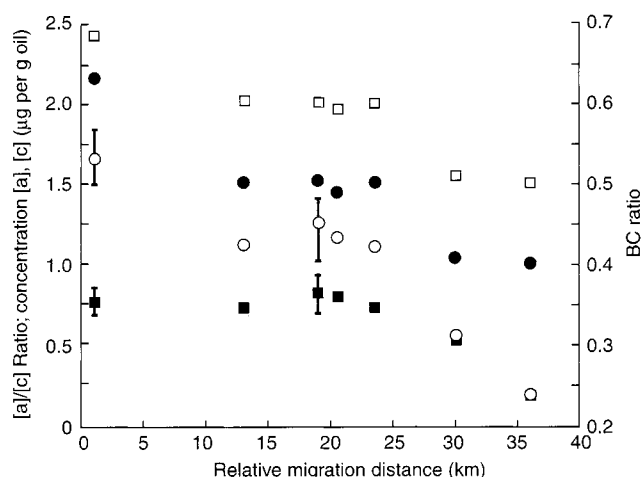


FIG. 4 Variation in the benzo[a]carbazole/(benzo[a]carbazole + benzo[c]carbazole) ratio (BC ratio, right axis; empty square symbols), benzo[a]carbazole/benzo[c]carbazole ratio ([a]/[c] ratio, left axis; filled circles) and concentrations of benzo[a]carbazole (empty circles) and benzo[c]carbazole (filled squares) for a suite of oils from the Central Graben area of the N. Sea (data set 3). Data range bars for samples at 1, 19 and 36 km show the maximum range of variation in up to five replicate analyses. The range bars on the sample at 36 km fall within the symbols.

oilfields are charged by long-range migration pathways rather than from more local source basins, or vice versa. In basins such as the North Sea, a major goal today is the discovery of small new oilfields either along migration pathways to the many known large accumulations, or on spill routes from one large oil accumulation to another. A simple ranking of discovered oils in terms of relative migration distance using benzocarbazole data appears useful in resolving the actual migration route from the potentially many connecting the accumulations, effectively defining areas

most likely to contain further small, but economically viable, accumulations.

The benzocarbazoles appear to have many ideal properties which contribute to their use as potential chemical odometers (milometers) for petroleum secondary migration. Investigation of other functionalized aromatic compounds with polarity/shape differences between isomers may lead to the discovery of other molecular parameters suitable for monitoring migration. Despite the widespread use of petroleum migration simulation programs in basin studies, current migration modelling efforts are often unconstrained by calibration data. Just as thermal maturation assessment has been used to successfully constrain basin thermal histories, a detailed quantitative understanding of the processes by which benzocarbazoles (or other petroleum components) are removed from petroleum during secondary migration may enable oil composition to be used as a quantitative constraint for computer models of the migration process. □

Methods

Analysis of benzocarbazoles. Data sets 1, 2 and 3. A pyrrolic-nitrogen enriched fraction was obtained from the de-asphalted oil by gravity column chromatography on alumina and on silicic acid²⁶. The benzocarbazoles were analysed quantitatively by gas-chromatography mass-spectrometry and selected ion mass spectrometry (GCMS-SIM) using a Hewlett Packard 5892 series mass selective detector and *N*-phenylcarbazole as internal standard¹⁰. Detection limits were 8 ng per g oil. Coefficients of variation determined from peak area measurements were $\leq 10\%$ for concentrations and $< 5\%$ for abundance ratios for oils migrated ≤ 100 km. **Data set 4.** A non-hydrocarbon (resin) fraction was obtained by medium pressure liquid chromatography on silica gel, and the benzocarbazoles were analysed by GCMS-SIM using a VG Autospec mass spectrometer operated at a resolution of 3,000 and *N*-phenylcarbazole as internal standard. This method has been found to give comparable results to those obtained using the analytical procedure described above (B.F.J.B., K.N. and A. Steen, unpublished results). **Data set 5.** A nitrogen-enriched fraction was obtained using a modification of a method described by Bennett, Bowler and Larter²⁷ for the analysis of phenols. The whole oil was applied to a non-encapped C18 SPE column and interferences were removed by elution with hexane; nitrogen compounds were then eluted with dichloromethane. The benzocarbazoles were quantified using GCMS-SIM as described for data sets 1, 2 and 3, above. These methods generate equivalent data.

Received 22 May; accepted 13 September 1996.

1. Magoon, L. B. & Dow, W. G. *The Petroleum System—From Source to Trap* (Am. Assoc. Petroleum Geologists, Tulsa, Mem. 60, 1994).
2. England, W. A. & Fleet, A. J. (eds) *Petroleum Migration* (Geological Soc., London, 1992).
3. Krooss, B. M., Brothers, L. & Engel, M. H. in *Petroleum Migration* (eds England, W. A. & Fleet, A. J.) 149–163 (Geological Soc., London, 1991).
4. Mackenzie, A. S., Patience, R. L., Maxwell, J. R., Vandenbrouke, M. & Durand, B. *Geochim. Cosmochim. Acta* **44**, 1709–1721 (1980).
5. Peters, K. E. & Moldovan, J. M. *The Biomarker Guide* (Prentice Hall, New Jersey, 1993).
6. Horstad, I., Larter, S. R. & Mills, N. in *The Geochemistry of Reservoirs* (eds England, W. A. & Cubitt, J.) 159–184 (Geological Soc., London, 1995).
7. Ballentyne, C. J., O'Nions, R. K. & Coleman, M. L. *Geochim. Cosmochim. Acta* **60**, 831–849 (1996).
8. Larter, S. R. & Aplin, A. C. in *The Geochemistry of Reservoirs* (eds England, W. A. & Cubitt, J.) 5–32 (Geological Soc., London, 1995).
9. Yamamoto, M. *Org. Geochem.* **19**, 389–402 (1992).
10. Li, M., Larter, S. R., Stoddart, D. & Bjorøy, M. *Anal. Chem.* **64**, 1337–1344 (1992).
11. Li, M., Larter, S. R., Stoddart, D. & Bjorøy, M. in *The Geochemistry of Reservoirs* (eds England, W. A. & Cubitt, J.) 103–123 (Geological Soc., London, 1995).
12. Li, M., Larter, S. R. & Frolov, Y. B. *J. High Res. Chromatogr.* **17**, 230–236 (1994).
13. Creaney, S. & Allan, J. in *Classic Petroleum Provinces* (ed. Brooks, J.) 189–202 (Geological Soc., London, 1990).
14. Wakefield, L. L. et al. in *The Petroleum Geology of Northwest Europe* (ed. Parker, J. R.) 459–468 (Geological Soc., London, 1993).
15. Brinca, D. thesis, Univ. Newcastle (1996).
16. Li, M., Larter, S. R. & Mei, B. *Org. Geochem.* **22**, 893–902 (1994).
17. Li, M., Yao, H., Stasiuk, L. D., Fowler, M. G. & Larter, S. R. *Org. Geochem.* (submitted).

18. McClure, N. M. & Brown, A. A. in *Giant Oil and Gas Fields of the Decade 1978–1988* (ed. Halbouty, M. T.) 307–322 (Am. Assoc. Petroleum Geologists, Tulsa, 1992).
19. Chen, M. thesis, Univ. Newcastle (1995).
20. Fedors, R. F. in *CRC Handbook of Solubility Parameters and other Cohesion Parameters* 2nd edn (ed. Barton, A. F. M.) 162 (CRC Press, Inc., Boca Raton, 1991).
21. Sandvik, E. I., Young, W. A. & Curry, D. J. *Org. Geochem.* **19**, 77–88 (1992).
22. Charlesworth, J. M. *Geochim. Cosmochim. Acta* **50**, 1431–1435 (1986).
23. Satter, A., Shum, Y. M., Adams, W. T. & Davis, L. A. *Soc. Petrol. Eng. J.* **20**, 129–138 (1980).
24. Macgregor, D. S. *Petrol. Geosci.* **2**, 197–218 (1996).
25. Clayton, C. J. & Hay, S. J. *Bull. Geol. Soc. Denmark* **41**, 12–33 (1994).
26. Later, D. W., Lee, M. L., Bartle, K. D., Kong, R. C. & Vassilarios, D. *Anal. Chem.* **53**, 1612–1620 (1981).
27. Bennett, B., Bowler, B. F. J. & Larter, S. R. *Anal. Chem.* (in the press).
28. Connolly, M. L. *Quantum Chemistry Program Exchange #429 as Implemented by Nemesis* (Oxford Molecular Ltd, Oxford, 1990).
29. Piggott, N. & Lines, M. D. in *Petroleum Migration* (eds England, W. A. & Fleet, A. J.) 207–226 (Geological Soc., London, 1991).

ACKNOWLEDGEMENTS. J. Allan, our colleague and friend, tragically died after submission of this Article. We thank the EU JOULE and THERMIE programs for support and our collaborators in these programmes for insight and comment: A. Aplin, B. Jones, A. Fleet, L. Schwark and K. Sorbie for useful discussions; A. Young for a helpful review; Imperial Oil Resources Ltd. (Canada), Norsk Hydro (Norway), Saga Petroleum (Norway), Shell (The Netherlands) and BP (UK) for providing samples and permission to publish; Y. Hall for preparing the manuscript, C. Jeans and B. Brown for preparing the artwork.

CORRESPONDENCE and requests for materials should be addressed to S.R.L. (e-mail: steve.larter@newcastle.ac.uk).

Structure of a replication-terminator protein complexed with DNA

Katsuhiko Kamada^{*†}, Takashi Horiuchi[‡], Katsufumi Ohsumi[‡], Nobuo Shimamoto^{*} & Kosuke Morikawa[†]

^{*} Department of Genetics, The Graduate University for Advanced Studies, and National Institute of Genetics, 1111, Yata, Mishima, Shizuoka 411, Japan

[†] Protein Engineering Research Institute (Biomolecular Engineering Research Institute)[§], 6-2-3, Furuedai, Suita, Osaka 565, Japan

[‡] Department of Molecular Biomechanics, The Graduate University for Advanced Studies, and National Institute for Basic Biology, 38, Nishigonaka, Myodaiji, Okazaki, Aichi 444, Japan

The crystal structure of the *Escherichia coli* replication-terminator protein (Tus) bound to terminus-site (*Ter*) DNA has been determined at 2.7 Å resolution. The Tus protein folds into a previously undescribed architecture divided into two domains by a central basic cleft. This cleft accommodates locally deformed B-form *Ter* DNA and makes extensive contacts with the major groove, mainly through two interdomain β-strands. The unusual structural features of this complex may explain how the replication fork is halted in only one direction.

In the *Escherichia coli* chromosome, bidirectional DNA replication is initiated from a single origin. Two replication forks, driven by the replication machinery, travel in opposite directions around the circular chromosome and terminate in a region opposite the origin (Fig. 1a). Two components dictate the arrest of the progression of replication (reviewed in refs 1 and 2). One is a set of six DNA sequences, designated as *Ter*, each containing a consensus element about 20-base-pairs (bp) long (Fig. 1b). The other is a trans-acting protein encoded by the *tus* (terminus utilization substance) gene. The *tus* protein (Tus; M_r 36K) specifically binds to each *Ter* site as a monomer³, forming a protein–DNA complex. This complex halts the passage of the replication fork in only one direction and permits passage by the other replication fork (Fig. 1a). Three *Ter* sites with the same polarity are distributed in one half of the terminus region, and the other three with the opposite polarity are in the other half. These sites are oriented so as to restrain DNA replication proceeding in the terminus-to-origin direction. Thus, this polar fork-blocking mechanism ensures that the terminus region works as a replication-fork trap. A similar system for replication termination is also present in *Bacillus subtilis*⁴. Moreover, polar fork-blocking sites have been found in the yeast, pea, frog and human genomes^{5–8}. The presence of these sites indicates that there might be an evolutionary advantage of the replication-arrest system.

The molecular basis of the DNA replication-fork arrest by the Tus–*Ter* complex is believed to be the inhibition of the replicative DnaB helicase^{9,10}, as the DNA polymerase III holoenzyme, which is a component of the replication machinery, lacks strand-displacing activity. Two mechanisms have been proposed for the arrest of the DNA replication fork. One model involves specific protein–protein interactions between a domain of Tus and some components of the replication machinery^{9,11–15}. These specific interactions inhibit the unwinding activity of DnaB, depending upon directions. The other model suggests a simple physical block against the progression of the replication fork. This model aptly explains why the unwinding of various helicases is halted by the Tus–*Ter* complex in only one direction^{10,16,17}. There is experimental evidence to support each model, and thus a conclusion awaits new information about the tertiary structure of the complex at an atomic resolution.

Here we report the crystal structure of the Tus protein bound to a 16-base *Ter*-site DNA fragment (Fig. 1c) at 2.7 Å resolution (Table 1). The structure of the Tus–*Ter* complex reveals a new

DNA-binding motif and clarifies the specific protein–DNA interactions. This crystallographic study suggests a plausible molecular mechanism for the polarity.

Features of the Tus–*Ter* complex

The Tus protein adopts a previously undescribed backbone structure. On the basis of the distance map, the structure is divided

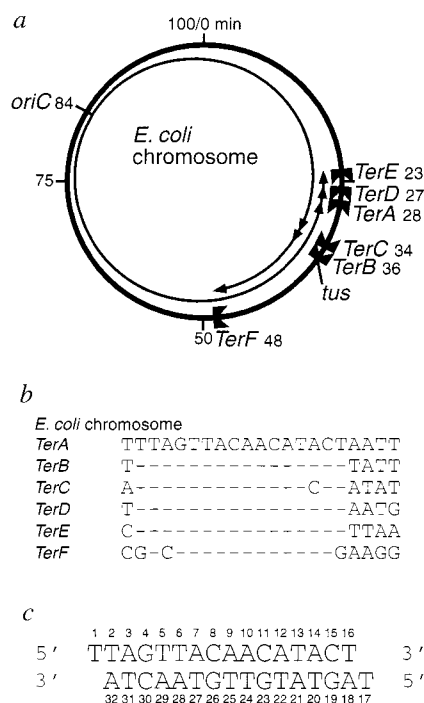


FIG. 1a, *Escherichia coli* chromosomal map representing the terminus region. The locations of the six terminus sites, replication origin (*oriC*) and the *tus* gene are shown. The symbols on the outer ring denotes the DNA-replication-terminus (*Ter*) sequence that blocks the replication fork approaching from the side towards the cut out arrowhead. Arrows represent bidirectional replication forks. The set *TerA*, *D* and *E* stops replication forks that travel anticlockwise on the chromosome. The other set, *TerC*, *B* and *F*, arrests clockwise-travelling forks. b, Comparison of the *Ter* sequences of *E. coli*. These sites arrest replication forks approaching from the 3' side (right). c, The synthetic *Ter* DNA duplex is 15-bp long with 5' overhanging thymines.

into two (amino and carboxy) domains (Fig. 2a, b). This overall structure has three distinct regions, two α -helical regions and central β structures, which jointly form a central large cleft. These prominent features are reminiscent of the pol domain of the Klenow fragment¹⁸, although the main-chain folding is different

in detail. The *Ter* DNA is accommodated into the positively charged central cleft, as the two protruding α -helical regions and the core β structures embrace 13 bp of the DNA duplex (Fig. 2a). The protein–DNA interface buries about 5,300 Å² of solvent-accessible area.

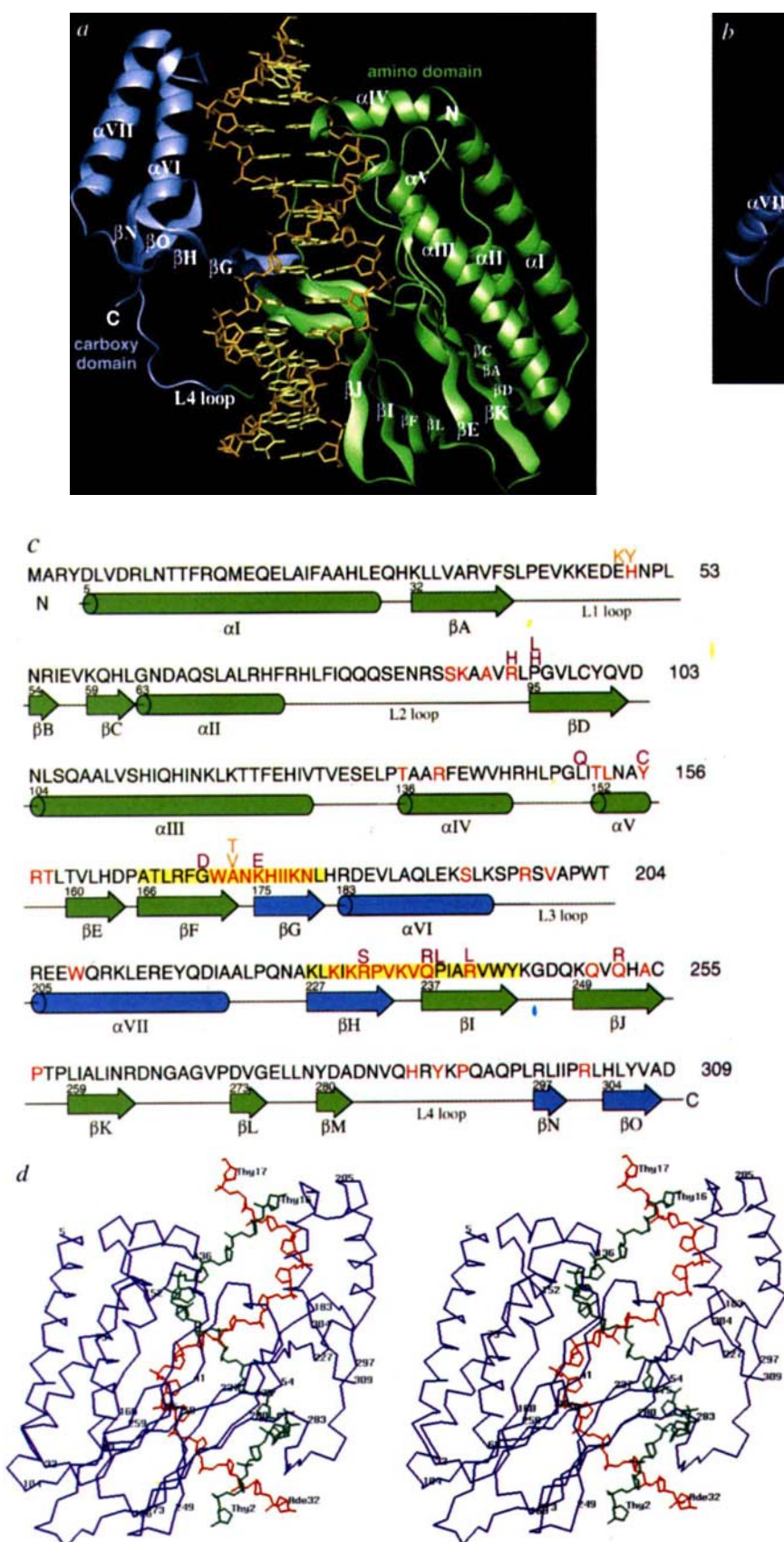


FIG. 2a, Ribbon drawing of the complex viewed perpendicular to the helical axis of the DNA. Amino and carboxy domains are coloured green and blue, respectively. b, Ribbon drawing of the complex displayed from the fork-blocking side. c, Sequence and secondary structure of Tus. The secondary structure elements are labelled and colour-coded as in a. Residues in the interdomain β -strands, which form the main elements for DNA recognition, are highlighted with yellow, and residues that interact with the DNA are indicated by red letters. Mutant residues are represented by violet and orange letters. d, C α stereo diagram of Tus with DNA backbones.

The two domains are connected mainly through well-defined, twisted, antiparallel β -strands (βF – βG and βH – βI) and the less-ordered, extended L4 loop (residues 283–296) (Fig. 2c). These interdomain β -strands, which are portions of two β -sheets of the two domains, contribute to the extensive contact with the bases of the major groove and the sugar-phosphate backbones of the *Ter* DNA. These interdomain β -strands lack main-chain hydrogen bonds at the central three residues (Asn 174, Lys 235, and Val 236) around the border of the two domains. These bulges and two proline residues (Pro 233 and Pro 238) allow the two β -strands to cross each other (Fig. 2d).

Tus protein. The amino and carboxy domains are both classified as $\alpha + \beta$ structures. The amino domain (residues 5–174 and 237–282) consists of an α -helical region, a β -sandwich, and three loops near the DNA. Three amphipathic α -helices (αI , αII , and αIII) form an antiparallel helix bundle aligned roughly in parallel with the DNA (Fig. 2b). Two αIV and αV helices, which are connected by a turn across a DNA backbone, clamp the phosphate backbones from both grooves (Fig. 2a, d). This α -helical region contributes to one-half of the amino-domain hydrophobic core. The remaining internal hydrophobic core is located within a β -sandwich (labelled $\beta CADKE/\beta LFIJ$, Fig. 2a, d). Its inner sheet ($\beta CADKE$) contacts the α -helical region. The other sheet ($\beta LFIJ$) is amphipathic and contributes to DNA binding. The βFI strands of the sheet form one-half of the interdomain β -strands (Fig. 2c). The last strand in this sheet (βL) extends into a short sheet (βBM) and subsequently turns towards the long L4 loop (Fig. 2d). This loop connecting the two domains is partly inserted into the minor groove. Two long loops with highly charged residues, L1 (41–53) and L2 (75–94), adopt extended conformations (Fig. 2b, d).

An internal peptide region (residues 175–234) and the C-terminal segment (residues 297–309) are jointly folded into the carboxy domain, which is stabilized by a hydrophobic core formed by α -helices (αVI and αVII) and a β -sheet ($\beta GHNO$, Fig. 2a). The βGH strands of the sheet extend to constitute the remaining

half of the interdomain β -strands (Fig. 2c). The L3 loop (residues 196–204) between the two α -helices contacts the minor groove of the DNA.

***Ter* DNA conformation.** A region of the *Ter* DNA duplex, ranging from base pairs T5·A29 to A9·T25, deviates significantly from the canonical B-form DNA, because of the binding to the interdomain region (Fig. 3a). The DNA duplex is underwound, with an average local helical twist of 29.5° and an average positive roll angle of 8.1° . Moreover, the neighbouring backbone from bases G4 to T6 is sandwiched between βF – βG of the interdomain β -strands and the L4 loop (Figs 2d and 3a). This DNA backbone region is deformed by their close contact. The torsion angle about the C4'–C5' bond of T5 assumes a *gauche-trans* conformation. The spacings between the contiguous phosphates from T5 to A7 deviate from the standard value of 6.7 Å in B-DNA. They result in 7.16 Å and 5.84 Å, with the T6 phosphate dragged towards A7. A T5·A29 base pair has the highly negative propeller angle of -24.2° . Consequently, the major groove, which accommodates the interdomain region, becomes deeper, and the minor groove is significantly expanded. These local deformations generate an overall DNA bend of about 20° .

Protein–DNA interactions

Backbone contacts. The Tus protein makes extensive polar contacts with 18 of the 26 phosphates along the embraced region of the duplex, though direct and indirect water-mediated hydrogen bonds (Fig. 3b). Notably, the interdomain region of Tus makes asymmetrical contacts with the phosphate backbones on the side, which allows passage of the fork. In fact, the phosphate backbone of A3–A7 forms many polar interactions with this region, whereas the complementary strand, T27–T31, is mostly exposed to the solvent (Fig. 3b). These features agree well with the results obtained by a DNA footprint analysis¹⁹.

The replication fork is stalled on the side of the two protruding α -helical regions from both domains (Fig. 2b). The αIV and αV

TABLE 1 Summary of crystallographic data

Diffraction data (<i>I</i> > 1σ)				MIR analysis (50–2.7 Å)					Figure of merit acentric/centric
Data set	Resolution (Å)	Unique refraction	Completeness (%)	<i>R</i> _{sym} (%)	<i>R</i> _{den} (%)	Number of sites	Phasing power acentric/centric	<i>R</i> _{culls} (%)	
Native	2.7	14,018	90.4	6.0					0.752/0.882
Pt1	3.0	7,386	63.0	8.1	30.7	4	0.88/0.52	0.86	
Pt2	2.7	13,727	86.4	4.9	11.1	4	0.50/0.33	0.88	
Hg1	3.0	9,770	83.6	6.5	30.2	2	1.61/0.85	0.76	
Hg2	2.7	14,023	89.7	4.5	16.3	2	1.01/0.53	0.84	
Hg3	2.7	13,273	86.4	6.9	30.6	2	1.36/0.78	0.81	
IdU-2-31	3.0	10,390	89.5	8.3	11.0	2	0.53/0.18	0.95	
IdU-2-27-31	2.7	10,055	79.5	5.9	13.1	3	1.55/0.54	0.77	
IdU-2-27	2.7	13,479	84.7	5.4	12.4	2	1.60/0.60	0.75	
IdU-2	2.7	13,940	88.4	5.2	6.5	1	0.54/0.22	0.94	
Refinement Statistics (10.0–2.7 Å)									
	No. of reflections		<i>R</i> factor (%)		<i>R</i> _{free} factor (%)				
Data with <i>I</i> > 2σ	12,726		19.6		31.4				
All data	14,517		22.0		32.6				
Total nonhydrogen atoms	3,103								
Water molecules	49								
Rms deviations from ideality					Protein		DNA		
Bond length (Å)					0.012		0.025		
Bond angle (degrees)					1.70		2.67		
Mean <i>B</i> value for main-chain atoms (Å ²) (r.m.s.)					25.71 (2.61)				
Mean <i>B</i> value for side-chain and DNA atoms (Å ²) (r.m.s.)					29.87 (3.57)		33.05 (5.73)		

$R_{\text{sym}} = \sum |I - \langle I \rangle| / \sum I$, where I is the observed intensity ($I > 1\sigma$), $\langle I \rangle$ is the average intensity of multiple observations of symmetry-related reflections.
 $R_{\text{den}} = \sum ||F_{\text{PH}}| - |F_{\text{P}}|| / \sum |F_{\text{P}}|$, where $|F_{\text{P}}|$ is the protein structure factor amplitude and $|F_{\text{PH}}|$ is the heavy-atom derivative structure factor amplitude.
 Phasing power = r.m.s. ($|F_{\text{H}}|/E$), where $|F_{\text{H}}|$ is the heavy-atom structure factor amplitude and E is the residual lack of closure error.
 $R_{\text{cullis}} = \sum |E| / \sum ||F_{\text{PH}}| - |F_{\text{P}}||$.
 Mean figure of merit = $\langle |\sum P(\alpha)e^{i\alpha} / \sum P(\alpha)| \rangle$, in which α is the phase and $P(\alpha)$ is the phase probability distribution.
 $R = \sum ||F_{\text{P}}| - |F_{\text{P(calc)}}|| / \sum |F_{\text{P}}|$.

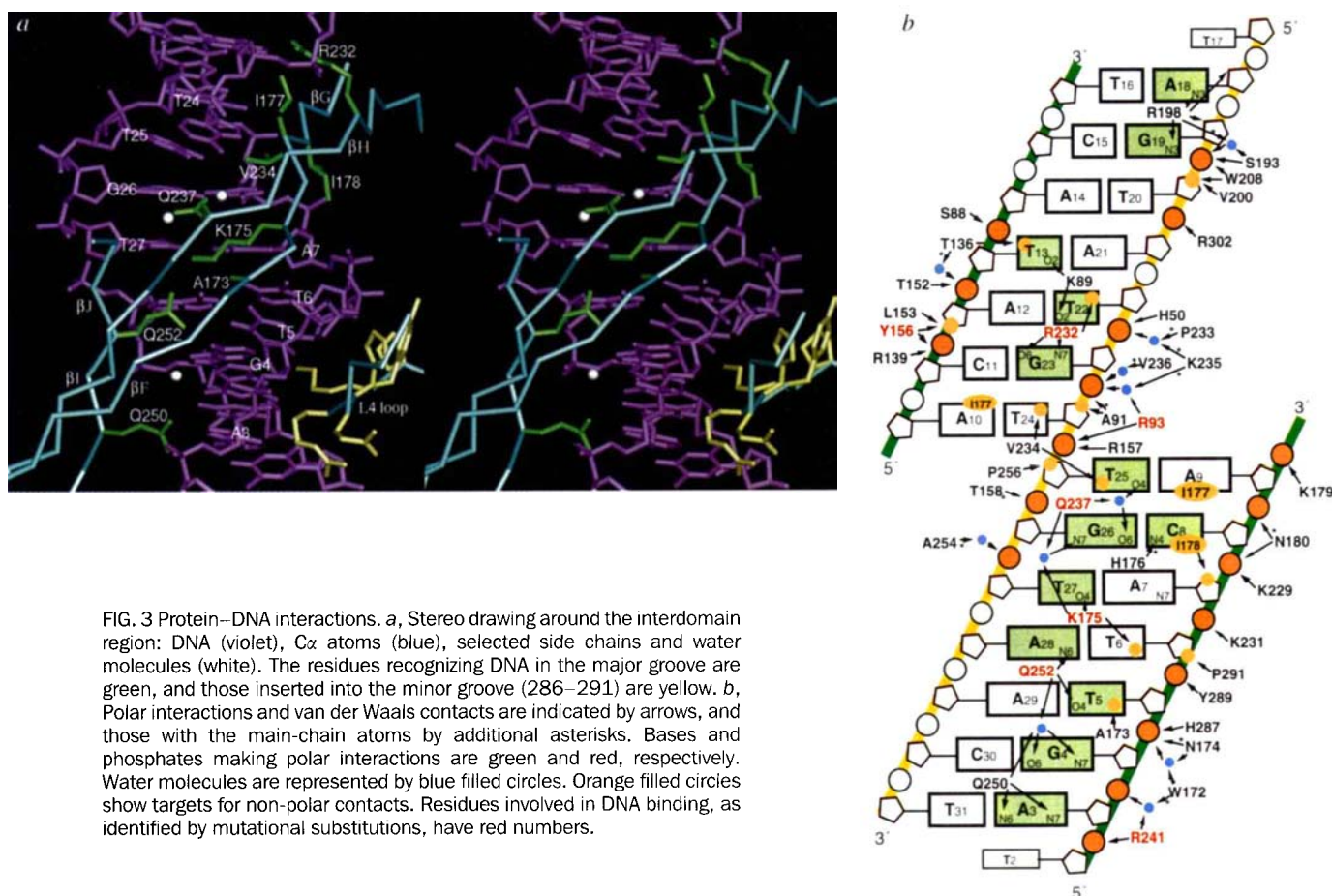


FIG. 3 Protein-DNA interactions. *a*, Stereo drawing around the interdomain region: DNA (violet), $C\alpha$ atoms (blue), selected side chains and water molecules (white). The residues recognizing DNA in the major groove are green, and those inserted into the minor groove (286–291) are yellow. *b*, Polar interactions and van der Waals contacts are indicated by arrows, and those with the main-chain atoms by additional asterisks. Bases and phosphates making polar interactions are green and red, respectively. Water molecules are represented by blue filled circles. Orange filled circles show targets for non-polar contacts. Residues involved in DNA binding, as identified by mutational substitutions, have red numbers.

helices and the L3 loop in these regions make several contacts with the sugar-phosphate backbones of both strands through polar and van der Waals interactions (Fig. 3*b*). The contact area on this side seems to be less extensive than that on the passage side, in spite of the large size of the α -helical regions.

Base recognition. The two β -sheets (β LFIJ and β GHNO) connected by the interdomain β -strands are important in the recognition of the *Ter* sequence by partial insertion into the major groove. In the amino domain, Gln 250 and Gln 252, which lie in the twisted β J strand, deeply penetrate into the groove to make contact with the bases (Fig. 3*a*). The side chain of Gln 250 forms bidentate hydrogen bonds with base A3, and the side chain of Gln 252 makes hydrogen bonds to the A28 and T5 bases. Both Gln residues also form water-mediated hydrogen bonds to the G4 base (Fig. 3*b*).

In the middle of the interdomain β -strands, the side chain of Gln 237 forms two water-mediated hydrogen bonds with bases T25 and G26 (Fig. 3*b*). The amino group of Lys 175 hydrogen bonds with base T27 and also forms a hydrogen bond with the side chain of Gln 252, located in the β J strand (Fig. 3*a*). This latter bridge-like contact contributes to the formation of a large hydrogen-bond network responsible for sequence recognition and stabilizes the twisted architecture of the interdomain β -strands in the major groove. Furthermore, the guanidium group of Arg 232 in the β H of the carboxy domain forms bidentate hydrogen bonds to the G23 base. On the side on which the fork is blocked, the L2 and L3 loops, to a smaller extent, recognize bases in the minor groove through the side chains of Lys 89 and Arg 198, respectively (Fig. 3*b*).

In addition to these polar interactions, van der Waals and hydrophobic contacts are important in DNA recognition. Particularly, in the carboxy domain, the majority of the contacts by the β -sheet seem to be non-polar. The side chains of Ile 177 and

Ile 178 jointly make van der Waals contacts with the sugar moiety of A7 and the bases of C8, A9 and A10 (Fig. 3*a*). The Arg 232 and Val 234 side chains form hydrophobic contacts with the methyl groups of the T22 and T25 bases, respectively (Fig. 3*b*). These are consistent with the results that the substitution of uracil for these bases reduces the stability of the complex²⁰. In the centre of the interdomain β -strands, the side chain of Ala 173 and the methylene group of Lys 175 make a hydrophobic interaction with the methyl groups of base T5 and T6, respectively (Fig. 3*b*).

Comparison with β -sheet motifs. DNA recognition motifs with β -sheets are found in some transcriptional regulatory proteins, such as MetJ, Arc, and TBP^{21–23}. In the case of MetJ and Arc, an intersubunit two-stranded antiparallel β -sheet is inserted into the major groove of a B-DNA duplex, in parallel with the phosphate backbones. TBP, with a curved antiparallel β -sheet, binds to the widened minor groove of TATA-element DNA. In these cases, the DNA bases are recognized by the residues on one side of the sheet. In the Tus-*Ter* complex, the two recognition sheets, connected by the interdomain β -strands, are almost perpendicularly inserted into the major groove (Fig. 2*d*). Notably, in the carboxy domain, the side chains on both sides of the β GH sheet recognize bases and phosphates (Fig. 3*a*). Thus, the Tus-*Ter* complex involves a previously undescribed DNA-binding scheme.

Tus mutants

We have screened Tus mutants that alter the ability to arrest DNA replication. Twelve types of mutant containing single-nucleotide substitutions were isolated (Figs 2*c* and 4; violet coloured letters and residues), and all of these mutants exhibit either a partial or complete loss of DNA-binding activity, as revealed by a gel-shift assay (data not shown).

Most of the mutations are mapped to the interdomain region

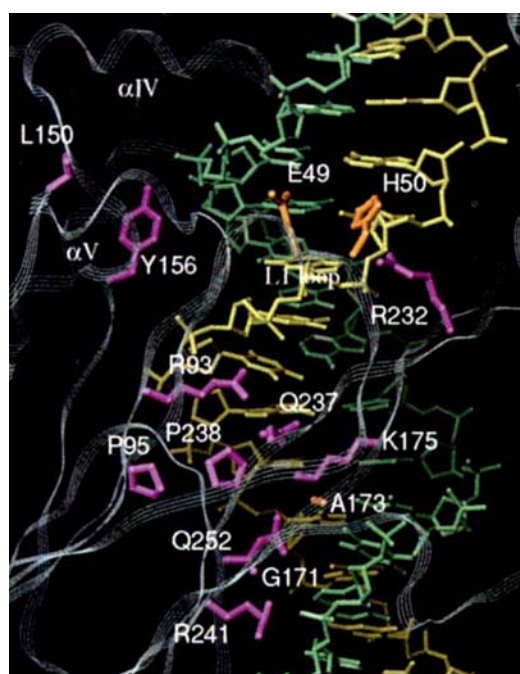


FIG. 4 Mutation sites mapped in the complex structure. The substitutions of violet and orange residues caused deficiencies in the replication-fork arrest. The two colours correspond to the single mutation sites shown in Fig. 2c.

involved in DNA binding (Fig. 4). The K175E, R232S, Q237R and Q252R mutations may disturb the hydrogen-bond network responsible for the sequence recognition in the major groove (Fig. 3b). The structure also reveals that, near the interdomain β -strands of the amino domain, Arg 93, Pro 95 and Pro 238 make hydrophobic contacts with each other. Thus, mutations R93H, P95H, P95L and P238L might destabilize the hydrophobic environment and prevent binding to *Ter* DNA.

The L150Q mutation would destabilize the hydrophobic region between the α IV and α V helices in the amino domain. The side chain of Tyr 156 hydrogen bonds to a phosphate of A12 and partly contributes to the hydrophobic region. The Y156C mutation would disrupt the hydrogen bond and disturb the hydrophobic region. These mutational observations suggest that the orientation of the α IV and α V helices is important for the replication arrest.

Several mutants have been isolated previously with defects in *in vivo* fork-blocking activity, and some of the mutants have been characterized. (A173V/T, E49K and H50Y)^{14,15} (Fig. 2c and 4, orange coloured letters and residues). The complex structure indicates that the A173V/T mutation may lower the fork-blocking activity, as the replacement of Ala by the bulky residues would inhibit DNA binding. The other two mutation sites are located in the flexible L1 loop, which contacts to DNA only through His 50. The E49K and H50Y mutations could result in conformational changes of the L1 loop, which affect DNA binding. Hence, it is likely that these mutants retain significant but considerably lower fork-blocking activity than the wild type.

Comparison with *Bacillus subtilis*

The *Bacillus subtilis* replication-terminator protein^{1,2} (RTP) is strikingly different from Tus, in terms of its biochemical properties and primary sequence, although it arrests replication in a similar polar manner. For instance, RTP forms a dimer in solution, and two dimers bind to each terminator (*TerI–TerVI*)²⁴. The *B. subtilis* *Ter* sequences are also very different from those of *E. coli* *Ter*, irrespective of their similar arrangement in each bacterial chromosome²⁴. The crystal structure of RTP, which recently has

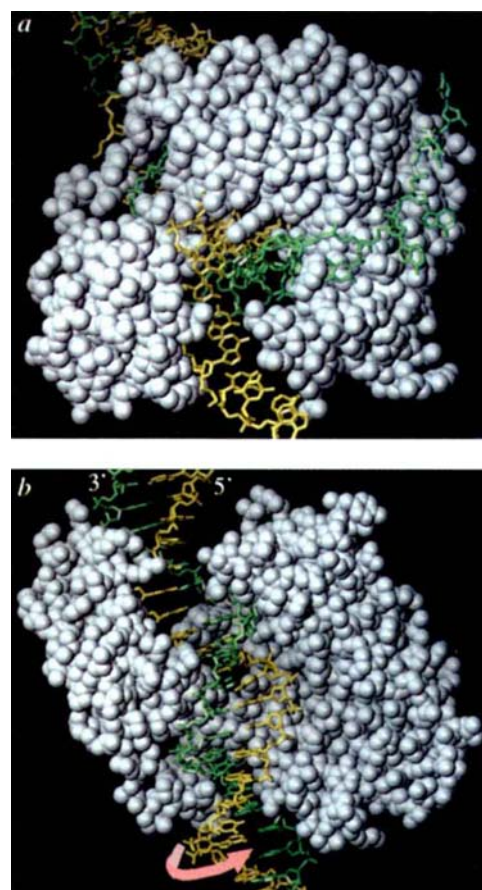


FIG. 5 a, Tus complexed with a putative Y-forked DNA, as viewed from the fork-block side. The single stranded DNAs near the terminus were hypothetically constructed to realize the inhibition of the fork progression. The two DNA strands are colour-coded as in Fig. 3a. b, A view from the free-passage side. Arrow (pink) denotes the putative unwinding motion of the DNA. The unwinding of the helicase from the lower right would displace the yellow DNA strand.

been determined in a DNA-free form is also different from that of Tus, in terms of the main-chain folding and the proposed binding motif^{25,26}. Furthermore, the *B. subtilis* RTP–*Ter* system efficiently arrests the progression of a replication fork in *E. coli*^{27,28}. This finding indicates that the fork arrest by the RTP–*Ter* complex in *B. subtilis* does not depend on the specific recognition and interaction between the RTP and a component of the approaching replication machinery.

From an evolutionary viewpoint, *E. coli* and *B. subtilis* might have independently developed replication-arrest systems consisting of two components, a terminator sequence and its specific binding protein. This emphasizes the advantage of such two-functional units for the directional replication fork trap, but no requirement of conserved tertiary structures of the terminator proteins.

Implications for directional arrest

The complex structure implies that the Tus protein tightly binds the DNA mainly through the interdomain β -strands, as if it were gripping the DNA duplex. The DNA duplex could not be released without a large conformational change of the Tus protein. The gel-shift assay also indicated that all the mutations reduced the ability of Tus to bind to the *Ter* sequence (data not shown), suggesting that the high affinity of the wild-type Tus for the *Ter* DNA ($K_d = 3.4 \times 10^{-13}$ M)¹⁹, is required for the fork arrest. This remarkably strong binding may explain the extensive interactions between Tus and the *Ter* DNA.

The unique structural features of the complex suggests to us a

hypothetical model for replication arrest. We have shown that α -helices are concentrated on the fork-blocking side. The two protruding α -helical regions from both domains may clasp the DNA duplex so that the interdomain β -structure, which is most important in DNA recognition, could be protected from direct contacts with unwinding proteins like DnaB helicase. Furthermore, the arrangements of the α IV and α V helices might effectively entangle a single-strand DNA, which the DnaB helicase would strip off (Fig. 5a, green strand). Consequently, the fork progression would be abolished. Such a process would involve very tight binding of Tus to the *Ter* site and steric impediment to helicase progression. Indeed, this model explains the reason why various helicases including DnaB are halted by the complex in only one direction^{10,16,17}.

On approach from the other side, the steric impediment is absent and the unwinding motion of helicases could relax the hold of Tus onto the DNA. The polar interactions of phosphate backbones with the interdomain β -structure are strikingly biased to one strand. This asymmetrical recognition could allow the DnaB helicase to strip off the other exposed strand more easily (Fig. 5b, yellow strand), accompanying the disruption of the primary binding region. As a result, the replication fork would pass the *Ter* site. Thus, the intrinsic structure of the complex may account for the polarity of the replication fork.

In conclusion, the three-dimensional structure of the complex explains at a molecular level the main aspects of the Tus-*Ter* protein-DNA interactions. This structural study provides the first indication of how replication forks are blocked or allowed to pass through the *Ter* site. □

Methods

Crystallographic analysis. Crystals for data collection were obtained from a mixture of the Tus protein with a 16-oligonucleotide DNA duplex containing the *Ter* consensus element²⁹. They belong to the space group $P4_12_12$, with unit cell parameters $a = 68.15$, $c = 230.72$ Å, and contain one protein bound to the DNA duplex per asymmetrical unit. The crystals diffract to 2.7 Å resolution on the a - b plane at room temperature, whereas reflections are observed beyond 2.0 Å along the c -axis. Intensity data of native and heavy-atom derivative crystals were collected with a Weissenberg camera³⁰ at the KEK Photon Factory on synchrotron beamline BL6A2 ($\lambda = 1.00$). Some data sets for derivatives were collected with a DIP100S imaging plate system (MAC science). All of the intensity data were processed using the program MacDENZO³¹. Heavy-atom search was carried out by conventional soaking methods and by co-crystallizations with *Ter* DNA containing 5-iododeoxyuridines (IdU). The mercury and platinum

derivatives were prepared by soaking native crystals in dialysis buffer including 15% (w/v) PEG4000 and one of the following heavy-atom compounds: 2.5 mM K_2PtCl_4 for 20 h (Pt1), 1 mM K_2PtCl_4 for 30 h (Pt2), 5 mM CH_3HgCl for 4 h (Hg1 and Hg3), and 2.5 mM CH_3HgCl for 4 h (Hg2). IdU derivatives were prepared by substituting 5-iodouracil for thymine at specific positions (residues 2, 27, 31) in the *Ter* DNA (Fig. 1c). Initial phases were obtained from two major Hg sites of Hg1 derivative identified using the Patterson search program HASSP^{32,33}, and the Pt and I sites were revealed by difference Fourier syntheses. Multiple isomorphous replacement and anomalous scattering (MIRAS) phases at 2.7 Å resolution generated from all derivatives were calculated by the use of MLPHARE^{33,34}. The initial molecular envelope (50% solvent content) was calculated automatically from the MIRAS map and then was subjected to 10 cycles of solvent flattening using programs from the CCP4 group³³. The map showed fine electron densities of the protein and DNA molecules with good connectivities. The atomic model constructed using the graphics program O³⁵ was refined by simulated annealing with the program X-PLOR³⁶. To monitor the progress of the refinement, subsets of 5% data within each resolution range were partitioned and used to calculate an R_{free} value³⁷. The first refined model showed an R -factor of 0.260 ($R_{free} = 0.377$) for the intensity data between 10.0 and 3.0 Å. The final refinement, including water molecules and optimizing the individual temperature factors, provided the model with an R -factor of 0.196 ($R_{free} = 0.314$) for the intensity data between 10.0 and 2.7 Å. The refined model of the complex contains all the amino-acid residues from 5 to 309, the 30 nucleotide residues forming the DNA duplex, and 50 water molecules. All the protein residues were within the allowed regions of the Ramachandran plot. Some loop regions, such as residues 47–50, 81–86, 246–248 and 288–294, exhibited high-temperature factors (>70 Å²). The four N-terminal residues and the two nucleotides at the termini of the two respective strands (1 and 32) seem to be structurally disordered, because of the lack of interpretable electron densities in the final map. The accessible surface area, using a probe of 1.4 Å radius, and the DNA parameters were calculated with the QUANTA and RNA³⁸ programs, respectively. Molecular images were produced with the O and QUANTA programs.

Screening method for fork-arrest-deficient mutants. Parent strain (KHG1327) contains another *Ter* sequence inserted into a *leu* gene, so as to block movement of the fork in a clockwise direction. In this strain, the anticlockwise fork was terminated at authentic sites, but the clockwise-travelling fork was trapped at the inserted site. We named it the 'blocked' strain, as it had a region that was difficult to replicate³⁹. A *recA*⁻ derivative of the blocked strains did not grow, probably because of a lack of recombinational repair of the double-stranded breaks resulting from the replication-fork arrest at the *Ter* sites. Thus, the host strain (KHG1327) was constructed by the knockout of *tus* in a *recA*⁻ derivative. A plasmid (pHSG399-*tus*⁺)⁴⁰, which had undergone mutagenesis in a *mutD5* mutator strain (KH1269)⁴¹, was transformed into this strain. Because the presence of the plasmid harbouring *tus*⁺ was lethal for the host strain, only the transformants with a *tus*⁻ mutant plasmid grew. We collected 31 *tus*⁻ clones, from which the *tus*⁻ plasmids were extracted, and the mutated sites in the *tus* gene were determined by DNA sequencing. Of these mutants, 10 had nonsense mutations, 8 had frame-shift mutations and 13 had single base-substituted mutations. Q252R was isolated twice.

Received 5 July; accepted 17 September 1996.

- Hill, T. M. *Annu. Rev. Microbiol.* **46**, 603–633 (1992).
- Baker, T. A. *Cell* **80**, 521–524 (1995).
- Sista, P. R., Hutchison, C. A. & Bastia, D. *Genes Dev.* **5**, 74–82 (1990).
- Weiss, A. S. & Wake, R. G. *J. Mol. Biol.* **179**, 745–750 (1984).
- Brewer, B. J. & Fangman, W. L. *Cell* **55**, 637–643 (1988).
- Hernandez, P., Lamm, S. S., Bjerknes, C. A. & Van't Hof, J. *EMBO J.* **7**, 303–308 (1988).
- Little, R. D., Platt, T. H. K. & Schildkraut, C. L. *Mol. Cell Biol.* **13**, 6600–6613 (1993).
- Wiesendanger, B., Lucchini, R., Koller, T. & Sogo, J. M. *Nucleic Acids Res.* **22**, 5038–5046 (1994).
- Khatrl, G. S., MacAllister, T., Sista, P. R. & Bastia, D. *Cell* **59**, 667–674 (1989).
- Lee, E. H., Kornberg, A., Hidaka, M., Kobayashi, T. & Honuchi, T. *Proc. Natl Acad. Sci. USA* **86**, 9104–9108 (1989).
- Bedrosian, C. L. & Bastia, D. *Proc. Natl Acad. Sci. USA* **88**, 2618–2622 (1991).
- Hiasa, H. & Mariani, K. J. *J. Biol. Chem.* **267**, 11379–11385 (1992).
- Natarajan, S., Kaul, S., Miron, A. & Bastia, D. *Cell* **72**, 113–120 (1993).
- Skokotas, A., Wroblewski, M. & Hill, T. M. *J. Biol. Chem.* **269**, 20446–20445 (1994).
- Skokotas, A., Hiasa, H., Mariani, K. J., O'Donnell, L. & Hill, T. M. *J. Biol. Chem.* **270**, 30941–30948 (1995).
- Hidaka, M. et al. *J. Biol. Chem.* **267**, 5361–5365 (1992).
- Lee, E. H. & Kornberg, A. *J. Biol. Chem.* **267**, 8778–8784 (1992).
- Olis, D. L., Brick, P., Hamlin, R., Xuong, N. G. & Steitz, T. A. *Nature* **313**, 762–766 (1985).
- Gottlieb, P. A. et al. *J. Biol. Chem.* **267**, 7434–7443 (1992).
- Duggan, L. J. et al. *J. Biol. Chem.* **270**, 28049–28054 (1995).
- Somers, W. S. & Phillips, S. E. *Nature* **359**, 387–393 (1992).
- Raumann, B. E., Rould, M. A., Pabo, C. O. & Sauer, R. T. *Nature* **367**, 754–757 (1994).
- Nikolov, D. B. et al. *Nature* **360**, 40–46 (1992).
- Franks, A. H., Griffiths, A. A. & Wake, R. G. *Mol. Microbiol.* **17**, 13–23 (1995).
- Bussiere, D. E., Bastia, D. & White, S. W. *Cell* **80**, 651–660 (1995).
- Swindells, M. B. *Trends Biochem. Sci.* **20**, 300–302 (1995).
- Young, P. A. & Wake, R. G. *J. Mol. Biol.* **240**, 275–280 (1994).

- Kaul, S. et al. *Proc. Natl Acad. Sci. USA* **91**, 11143–11147 (1994).
- Kanada, K., Ohsumi, K., Horuchi, T., Shimamoto, N. & Morikawa, K. *Proteins* **24**, 402–403 (1996).
- Sakabe, N. *J. Appl. Crystallogr.* **16**, 542–547 (1983).
- Otinowski, Z. & Minor, W. in *The MacDenzo Manual* (ed. Gewirth, D.) (Yale Univ., New Haven, 1994).
- Terwilliger, T. C., Kim, S.-H. & Eisenberg, D. *Acta Crystallogr. A* **43**, 1–5 (1987).
- Collaborative Computational Project No. 4, SERC Daresbury Laboratory, Warrington, UK *Acta Crystallogr. D* **50**, 760–763 (1994).
- Otinowski, Z. in *Isomorphous Replacement and Anomalous Scattering* (eds Wolf, W., Evans, P. R. & Leslie, A. G. W.) 80–86 (SERC, Daresbury, 1991).
- Jones, T. A., Zhou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta Crystallogr. A* **47**, 110–119 (1991).
- Brünger, A. T. (ed.) *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, 1992).
- Brünger, A. T. *Nature* **355**, 472–475 (1992).
- Babcock, M. S., Pednault, E. P. D. & Olson, W. K. *J. Biomol. Struct. Dynam.* **11**, 597–628 (1993).
- Horiuchi, T. & Fujimura, Y. *J. Bacteriol.* **177**, 783–791 (1995).
- Takeshita, S., Sato, M., Toda, M., Masahashi, W. & Hashimoto-Goto, T. *Gene* **61**, 63–74 (1987).
- Maruyama, M., Horiuchi, T., Maki, H. & Sekiguchi, M. *J. Mol. Biol.* **167**, 757–771 (1983).

ACKNOWLEDGEMENTS. We thank M. Hidaka for contributions at the early stage of this study; S. E. Tsutakawa, H. Shinagawa and H. Iwasaki for helpful discussions; T. Hakoshima, L. Chapman and D. Rhodes for advice about DNA synthesis and purification; and M. Ariyoshi, N. Kunishima, D. G. Vassilyev, N. Sakabe, and N. Watanabe for assistance with data collection. We thank A. Kornberg for comments before submission.

CORRESPONDENCE and requests for materials should be addressed to K.M. (e-mail: morikawa@ns.berli.co.jp). Coordinates have been deposited in the Protein Data Bank, Brookhaven (accession number 1ECR).

Sodium abundance variations in main-sequence stars of the globular cluster 47 Tucanae

Michael M. Briley*, Verne V. Smith†‡, Nicholas B. Suntzeff§, David L. Lambert‡, Roger A. Bell|| & James E. Hesser¶

* Department of Physics and Astronomy, University of Wisconsin Oshkosh, Oshkosh, Wisconsin 54901, USA

† Department of Physics, University of Texas at El Paso, El Paso, Texas 79968, USA

‡ McDonald Observatory, University of Texas, Austin, Texas 78741, USA

§ National Optical Astronomy Observatories, Cerro Tololo Inter-American Observatory, Casilla 603, La Serena, Chile

|| Department of Astronomy, University of Maryland, College Park, Maryland 20742, USA

¶ Dominion Astrophysical Observatory, Herzberg Institute of Astrophysics, National Research Council, 5071 Saanich Road, RR5, Victoria, British Columbia V8X 4M6, Canada

GALACTIC globular clusters were once thought to be chemically homogeneous, having formed quite rapidly from relatively small condensations of primordial gas²⁹. In many clusters, significant star-to-star variations in light-element abundances have been observed¹⁻⁴ in evolved giant stars. These variations have been attributed to the presence at the stellar surfaces of nucleosynthesis products generated deep within the stars. But other observations¹³ have suggested that some of this variability was established earlier in the stars' lifetimes, perhaps as a result of inhomogeneities in the gas cloud from which the cluster formed. Here we report the observation of variations in the sodium abundances of unevolved (main-sequence) stars in the cluster 47 Tucanae. Although these variations are similar to those observed in evolved cluster stars, they cannot be explained by mixing, in the framework of current models of stellar evolution. This indicates either that the gas out of which 47 Tuc formed was chemically inhomogeneous, or that some mechanism for altering the surface element abundances of stars operates while they are still on the main sequence.

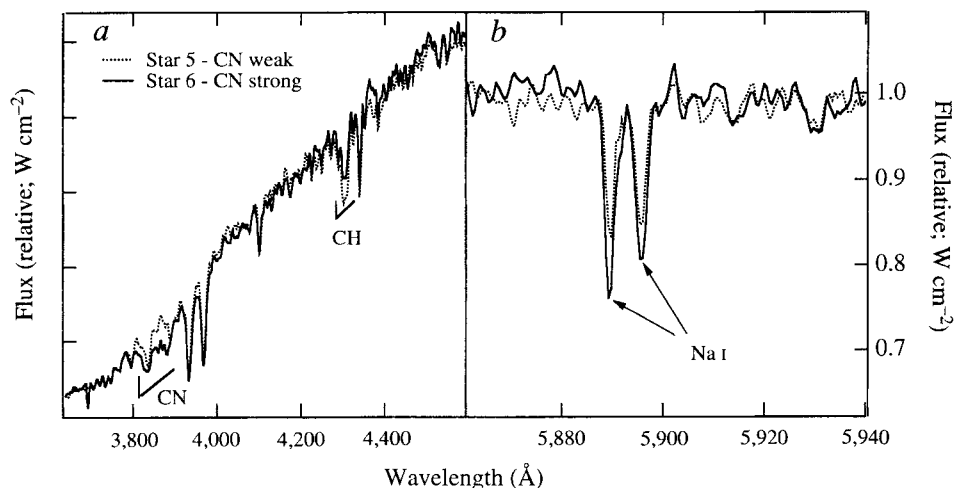
To distinguish effects due to internal stellar evolution on the red-giant branch (RGB) from those which may have been present before this point, we have observed unevolved, or main-sequence, stars in the nearby globular cluster 47 Tucanae. By virtue of its relatively high metallicity, $[Fe/H] = -0.8$, the main-sequence turn-off (MSTO) stars in 47 Tuc are cool enough for Na I lines to be detected and bright enough to be observed at moderate

spectral resolution. (Metallicity refers to the overall abundance of elements heavier than He. Here, $[A/B]$ denotes the logarithm of the ratio of the abundances of A to B with respect to the solar distribution.) The surface compositions of these main-sequence stars are widely supposed to reflect the chemical compositions of the cluster gas from which they formed. Among the brighter (evolved) 47 Tuc RGB stars, large variations in CN band strengths, as well as significant Na abundance variations, have been observed⁵⁻⁸. These observations are often cited as evidence for an episode of deep mixing during the first ascent of the RGB in which C- (and possibly O-) poor, N- and Na-rich material is brought to the surface from regions undergoing the CN (and possibly ON) cycles of H-burning and proton captures on Ne (refs 4, 9, 10). Suggested candidates driving this mixing include meridional circulation¹¹ and turbulent diffusion¹². Such processes, however, are currently not included in standard treatments of the structure of low-mass stars, and the extent to which either process operates in globular cluster stars remains unclear.

A group of 20 MSTO stars ($3.95 < M_V < 4.4$; M_V , absolute visual magnitude) with well determined photometry has already been observed in the blue CN and CH region and found to exhibit a large range in CN band strengths¹³ (Fig. 1a). In August 1995, we obtained spectra of the Na D lines (5,890, 5,895 Å) for a subset of these stars using the Cerro Tololo Interamerican Observatory 4-m telescope with the Ritchey-Chretien (RC) spectrograph. The programme stars were divided into three pairs of CN-strong and weak CN-stars of similar brightness and colours (Table 1). One pair was observed each night with alternating exposures between the CN-strong and CN-weak member. Of critical importance in the reduction procedure was the subtraction of the night-sky Na I emission features which were not appreciably shifted away from the stellar lines. This was accomplished by first removing image distortions perpendicular to the dispersion direction (by tracing the emission features) and then fitting the sky spectrum during the extraction process. Other regions of the spectra contaminated by sky emission—yet free of stellar lines—were also examined, to verify the subtraction process. Final spectra for stars 5 and 6 are plotted in Fig. 1b.

From the final summed spectra, equivalent widths were measured by direct numerical integration for the two Na I D-lines, as well as a Ca I line at 6,122 Å. The inherent error in the final widths was estimated by computing the width of each line in each of the individually measured spectra and taking the standard deviation about the mean value. The resulting widths and their associated uncertainties are listed in Table 1. The signal-to-noise ratio (S/N) of each spectrum was also measured in the region around the lines using a standard analysis package in the Image Reduction and Analysis Facility (IRAF, a software package available from the National Optical Astronomy Observatories) (Table 1). The

FIG. 1 a, b, Spectra of two 47 Tucanae MSTO stars of identical brightness and colour ($M_V = 3.95$, $(B - V)_0 = 0.55$; ref. 13) which therefore are also of the same temperature and surface gravity. a, The 3,883-Å CN band of star 5 appears considerably weaker than that of star 6. The strengths of the CN band at 4,350 Å are also anticorrelated with CN (see ref. 13). b, Final spectra of Na I lines from the two stars. The RC spectrograph and Loral 3K CCD were used with the no. 450 632 lines per mm grating in second order for a final resolution of 1.4 Å full-width at half-maximum. Each spectrum is the sum of four 3,000-s exposures taken over the nights of 17–19 August 1995. The reduction of the images followed standard procedures using the IRAF package.



expected error in the widths may be estimated from the relation¹⁴ $\delta W_{\text{eq}} \sim 1.6(wd)^{0.5} (S/N)^{-1}$, where δW_{eq} is the error in equivalent width, w is the line width, and d is the pixel size. For the Ca I lines, $d = 0.7 \text{ \AA}$, $w \approx 1.5 \text{ \AA}$ and $S/N \approx 50$, for $\delta W_{\text{eq}} \approx 30 \text{ m\AA}$. In the case of the wider Na I lines, $w \approx 2 \text{ \AA}$ and $S/N \approx 50$, with $\delta W_{\text{eq}} \approx 40 \text{ m\AA}$. Both results compare well to the estimates in Table 1.

In Fig. 2 we plot the Na I and Ca I equivalent widths versus the CN band strength excesses ($\delta S(3839)$) taken from ref. 13. In this figure, the Na I equivalent widths increase with $\delta S(3839)$ (that is, CN band strength), while the Ca I line shows no correlation with CN. Broken lines show linear least-squares fits and demonstrate a significant correlation ($r = 0.89$) of Na I line strength with $\delta S(3839)$, but no correlation ($r = -0.12$) of Ca I with CN. From both Figs 1 and 2, and Table 1, it is clear that significant variations in the Na abundances are present among these MSTO stars at the 2.75σ level (the mean difference in Na I widths between the CN-strong and -weak groups is $220 \text{ m\AA}$ with a mean error in a single measurement of $\pm 80 \text{ m\AA}$).

Also shown in Fig. 2 are arrowed lines which illustrate quantitative abundances representative of various Ca I and Na I equivalent widths. These abundances were calculated using a model atmosphere which is typical of the observed stars ($T_{\text{eff}} = 5,900 \text{ K}$, $\log g = 4.15$, microturbulence of 1.5 km s^{-1} , and $[\text{Fe}/\text{H}] = -0.8$) and the local-thermodynamic-equilibrium spectral analysis program MOOG¹⁵. The lines are labelled for abundances on a scale of $\varepsilon(X) = \log(N(X)/N(\text{H})) + 12$ (where $N(X)$ is the number abundance of atom X , normalized to $N(\text{H}) = 10^{12}$). As is apparent here, the Ca abundances show a range of ~ 0.2 dex (logarithmic ratio) around mean abundance of $\varepsilon(\text{Ca}) = 5.9$ ($[\text{Ca}/\text{Fe}] = +0.4$) with a 1σ dispersion of 0.1 dex. This range is well within the 1σ uncertainty of any one star (~ 0.4 dex). The Na abundances vary from $\varepsilon(\text{Na}) = 5.4$ to 5.8 ($[\text{Na}/\text{Fe}] = -0.1$ to $+0.3$, with a mean of $\varepsilon(\text{Na}) = 5.6$ ($[\text{Na}/\text{Fe}] = +0.1$) and a dispersion of 0.2 dex. Unlike the case of Ca I, the range in Na abundances (0.4 dex) is greater than the mean estimated 1σ error in any one Na measurement (~ 0.2 dex). Note that all six programme stars have quite similar temperatures and surface gravities, so it is unlikely that the Na I variations could be caused by peculiarities in the stars' atmospheres which might change Na I, but leave the rest of the stellar spectrum unchanged.

The present Na abundance patterns on the 47 Tuc main-sequence stars are very similar to those observed among evolved RGB stars in this cluster. In a classic study of four 47 Tuc bright red giants, Cottrell and Da Costa⁶ reported an $[\text{Na}/\text{Fe}]$ enhancement of 0.3 – 0.4 dex among the CN-strong as compared to the CN-weak stars (with an uncertainty of ± 0.1). More recent studies^{7,8} have used high-resolution spectroscopy of weaker Na lines in 47 Tuc giants. Their combined data set consists of seven stars which exhibit a range in $[\text{Na}/\text{Fe}]$ from -0.12 to $+0.28$, with a mean of $[\text{Na}/\text{Fe}] = +0.10$ and a dispersion of 0.14 dex. Using David Dunlap Observatory (DDO) photometry¹⁶, the CN band strengths of all four of the stars in ref. 8 may be classified as follows: CN strong, $[\text{Na}/\text{Fe}] = +0.19 \pm 0.09$ (star 1513), $+0.28 \pm 0.04$ (3501); CN weak, -0.01 ± 0.06 (2525), -0.12 ± 0.15 (4418). One star (8416) in ref. 7 has a known CN band strength (weak) with a reported $[\text{Na}/\text{Fe}]$ of $+0.11 \pm 0.13$.

The sample sizes involved are admittedly small and comparisons are risky, yet these results suggest that the abundances of Na among the MSTO and the evolved stars exhibit similar values. The implication is that little modification of surface Na abundances has taken place during the evolution of 47 Tuc stars from $M_V \sim +4.2$ to ~ -1.5 . The observed Na variations also appear

TABLE 1 Programme stars, colours, band strengths and equivalent widths

Star*	M_V	$W_{\text{eq}}(5,890 \text{ \AA})^\dagger$	$W_{\text{eq}}(5,895 \text{ \AA})^\dagger$	$\sum W_{\text{eq}} \text{ Na I}^\ddagger$	$W_{\text{eq}} \text{ Ca I}^\ddagger$
CN-strong stars:					
6	3.95	577 ± 24	553 ± 49	$1,130 \pm 65(57)$	$169 \pm 22(64)$
10	4.08	490 ± 66	468 ± 42	$958 \pm 103(47)$	$144 \pm 20(57)$
21	4.36	536 ± 51	344 ± 56	$880 \pm 114(44)$	$145 \pm 22(52)$
CN-weak stars:					
5	3.95	363 ± 29	388 ± 19	$751 \pm 41(55)$	$167 \pm 11(68)$
9	4.07	438 ± 40	292 ± 23	$730 \pm 33(60)$	$170 \pm 33(53)$
23	4.40	442 ± 50	383 ± 94	$825 \pm 113(41)$	$146 \pm 36(51)$

* Data from ref. 13.

† Equivalent widths measured in mÅ.

‡ Numbers in parentheses indicate the measured signal to noise ratio of the spectrum in the region of interest.

to be largely the result of a process which took place before the stars left the main sequence. The correlation of MSTO stars' Na abundances with CN band strengths further implies that whatever astrophysical process produced the excess N responsible for the CN-strong dwarf stars, also generated large amounts of Na.

At first glance, the present observations seem inconsistent with the decrease in C abundances with increasing luminosities seen among RGB stars in the most metal-poor clusters^{17–19}—which are widely accepted as evidence for some mixing process in those halo clusters. Observations of remarkable negative correlations between O and Na abundances among evolved bright red giants in other clusters have also been interpreted as the mixing of fresh Na to the surface^{20–22}. However, to the extent that our rudimentary theories of evolution-driven mixing and nuclear reaction networks

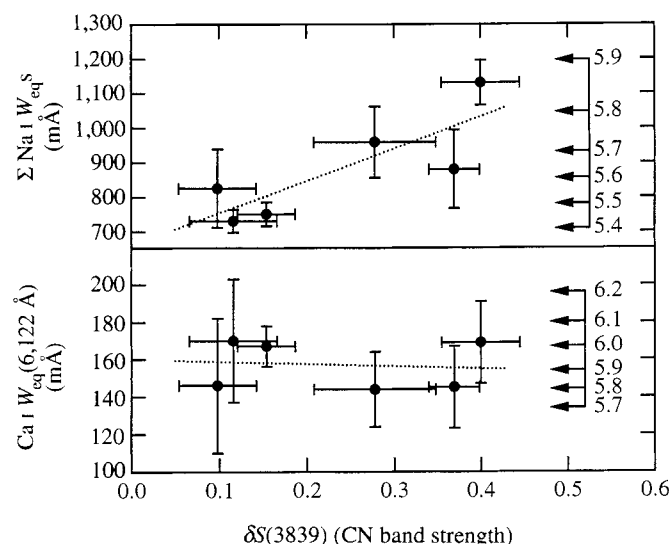


FIG. 2 Measured equivalent widths (line strengths) for the MSTO (unevolved) stars of Table 1, plotted against measures of their CN band strength ($\delta S(3839)$)¹³. The dashed lines are a linear least-squares fit to the data. Although there is no relation between CN band strength and Ca, the Na versus CN correlation is clearly apparent. The numbers to the right are the abundances corresponding to the indicated equivalent widths. The spread in Na line strengths corresponds to a difference of 0.4 in $[\text{Na}/\text{Fe}]$, with the CN-strong stars also having the higher Na abundances. The pattern of abundances is very similar to those that have been observed among the more evolved cluster members—which has recently been suggested (in the context of other clusters) to be the result of deep mixing. The present observations imply that a large component of these chemical inhomogeneities are present in the stars before they begin to evolve off the main sequence.

are applicable, these processes should be less efficient in more metal-rich stars such as those found in 47 Tuc^{10–12}. Observations of other clusters appear to confirm this^{23–25}. As has been suggested previously by numerous authors³, it is likely that both primordial (or 'pre-RGB' ascent) and mixing mechanisms have worked with varying degrees of efficiency in most clusters.

In the case of the disk globular cluster 47 Tuc, the present observations suggest that at least the bulk of the C, N and Na abundance variations were fixed before RGB ascent. One interpretation would be that the gas from which the 13–14-Gyr-old²⁶ 47 Tuc stars formed was more poorly mixed, in at least C, N and Na than was apparently the case for other clusters. Moreover, this history is different from that of the field stars, none of which are known to show Na enhancements²⁷. Candidates for this mechanism discussed in the literature include: thermal-pulse driven mass loss of triple-alpha processed material (as well as s-processed Na, Mg and Al) from the outer envelopes of 5–10 $M_{\odot}$ stars during an extended period of low-mass star formation⁶, and accretion of processed material ejected by intermediate-mass stars²⁸. A further

possibility lies in the main-sequence stars themselves: the present observations do not rule out some modification of surface C, N and Na abundances during the stars' 13–14 Gyr on the main sequence. Should this be the underlying cause, these stars must be more complex than our current models, which make no such prediction.

In either case, our understanding of the early environment and stellar evolution in the clusters which anchor our age dating of the Universe is far less than complete. Should these abundance variations be due to a primordial source, how could the 'polluting' material remain unmixed over the timescale required for low-mass star formation? What are the characteristics of the stars which injected this material (see, for example, Fig. 15 of ref. 28)? By continuing to disentangle the abundance signatures of the globular clusters through spectroscopy of their unevolved main-sequence stars and comparable field halo stars, we can hope to identify the events that shaped the chemistry of the early Galaxy, the physics of stellar evolution, and ultimately, to put cosmological age dating on a far firmer basis. □

Received 1 July; accepted 10 September 1996.

1. Smith, G. H. *Publ. Astron. Soc. Pacif.* **99**, 67–90 (1987).
2. Suntzeff, N. B. in *The Globular Cluster-Galaxy Connection* (eds Smith, G. H. & Brodie, J. P.) 167–179 (ASP Conf. Ser. Vol. 48, Astronomical Society of the Pacific, San Francisco, 1993).
3. Briley, M. M., Bell, R. A., Hesser, J. E. & Smith, G. H. *Can. J. Phys.* **72**, 772–781 (1994).
4. Kraft, R. P. *Publ. Astron. Soc. Pacif.* **106**, 553–565 (1994).
5. Norris, J. & Freeman, K. C. *Astrophys. J.* **230**, L179–L182 (1979).
6. Cottrell, P. L. & Da Costa, G. S. *Astrophys. J.* **245**, L79–L82 (1981).
7. Gratton, R. L., Quarta, M. L. & Ortolani, S. *Astron. Astrophys.* **169**, 208–218 (1986).
8. Brown, J. A. & Wallerstein, G. *Astron. J.* **104**, 1818–1830 (1992).
9. Langer, G. E. & Hoffman, R. *Publ. Astron. Soc. Pacif.* **107**, 1177–1182 (1995).
10. Cavallo, R. M., Sweigart, A. V. & Bell, R. A. *Astrophys. J.* **464**, L79–L82 (1996).
11. Sweigart, A. V. & Mengel, J. G. *Astrophys. J.* **229**, 624–641 (1979).
12. Charbonnel, C. *Astron. Astrophys.* **282**, 811–828 (1994).
13. Briley, M. M., Hesser, J. E., Bell, R. A., Bolte, M. & Smith, G. H. *Astron. J.* **108**, 2183–2192 (1994).
14. Cayrell, R. in *The Impact of Very High Signal to Noise Spectroscopy on Stellar Physics* (eds Cayrell de Strobel, G. & Spite, M.) 345–353 (Kluwer, Dordrecht, Boston, 1988).
15. Sneden, C. *Astrophys. J.* **184**, 839–849 (1973).
16. Norris, J. & Freeman, K. C. *Astrophys. J.* **230**, L179–L182 (1979).
17. Bell, R. A., Dickens, R. J. & Gustafsson, B. *Astrophys. J.* **229**, 604–623 (1979).

18. Carbon, D. F. et al. *Astrophys. J. Suppl. Ser.* **49**, 207–258 (1982).
19. Trefzger, D. V., Langer, G. E., Carbon, D. F., Suntzeff, N. B. & Kraft, R. P. *Astrophys. J.* **266**, 144–159 (1983).
20. Sneden, C., Kraft, R. P., Prosser, C. F. & Langer, G. E. *Astron. J.* **104**, 2121–2140 (1992).
21. Kraft, R. P., Sneden, C., Langer, G. E. & Shetrone, M. D. *Astron. J.* **106**, 1490–1507 (1993).
22. Pilachowski, C. A., Sneden, C., Kraft, R. P. & Langer, G. E. *Astron. J.* **112**, 545–564 (1996).
23. Bell, R. A. & Dickens, R. J. *Astrophys. J.* **242**, 657–672 (1980).
24. Suntzeff, N. B. & Smith, V. V. *Astrophys. J.* **381**, 160–172 (1991).
25. Briley, M. M., Smith, V. V. & Lambert, D. L. *Astrophys. J.* **424**, L119–L122 (1994).
26. Hesser, J. E., Shott, P., Harris, W. E., Allwright, J. W. B. & VandenBerg, D. A. *Publ. Astron. Soc. Pacif.* **99**, 739–808 (1987).
27. Pilachowski, C. A., Sneden, C. & Kraft, R. P. *Astron. J.* **111**, 1689–1704 (1996).
28. Norris, J. E. & Da Costa, C. S. *Astrophys. J.* **447**, 680–705 (1995).
29. Murray, S. D. & Lin, D. N. *Astrophys. J.* **339**, 933–942 (1989).

ACKNOWLEDGEMENTS. We thank the staff of the Cerro Tololo Interamerican Observatory, without whose assistance this project would not have been possible. The National Optical Astronomy Observatories are operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the US NSF. This work was supported in part by the Fund for Astrophysics, the UW Oshkosh Faculty Development Program, and NSF grants to M.M.B., V.V.S., D.L.L. and R.A.B.

CORRESPONDENCE should be addressed to M.M.B. (e-mail: mike@maxwell.phys.uwosh.edu).

Detection of acetylene in the infrared spectrum of comet Hyakutake

T. Y. Brooke*, A. T. Tokunaga†, H. A. Weaver‡, J. Crovisier§, D. Bockelée-Morvan§ & D. Crisp*

* Jet Propulsion Laboratory, M/S 169–237, 4800 Oak Grove Drive, Pasadena, California 91109, USA

† Institute for Astronomy, University of Hawaii, Honolulu, Hawaii 96822, USA

‡ Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, Maryland 21218, USA

§ Observatoire de Paris-Meudon, F-92195, Meudon, France

COMETS are rich in volatile materials, of which roughly 80% (by number) are water molecules¹. Considerable progress^{2–4} is being made in identifying the other volatile species, the abundances of which should enable us to determine whether comets formed primarily from ice-covered interstellar grains⁵, or from material that was chemically processed in the early solar nebula^{6,7}. Here we report the detection of acetylene (C_2H_2) in the infrared spectrum of comet C/1996 B2 (Hyakutake). The estimated abundance is 0.3–0.9%, relative to water, which is comparable to the predicted solid-phase abundance in cold interstellar clouds. This suggests that the volatiles in comet Hyakutake may have come from ice-

covered interstellar grains, rather than material processed in the accretion disk out of which the Solar System formed.

Our observations of Hyakutake were obtained at the 3-m NASA Infrared Telescope Facility atop Mauna Kea, Hawaii on 8 April 4:00–5:00 UT 1996 when the comet was at a heliocentric distance $r = 0.73$ AU, and a geocentric distance $\Delta = 0.47$ AU (1 AU is the Sun–Earth distance). The spectrum was obtained with CSHELL^{8,9}, an echelle spectrometer that provides a spectral resolution of 2×10^4 at $3.0 \mu m$ through its 1.0-arcsec-wide slit. The instrument records spatially resolved spectra on a 256×256 InSb array. For these observations, each pixel subtended 0.027 cm^{-1} in the dispersion direction and 0.2 arcsec in the spatial direction (70 km in projection at the comet). The 30-arcsec-long slit was oriented east–west and centred on the peak infrared signal from the comet, giving spatial coverage up to 5,000 km from the nucleus. The seeing estimated from stellar images was ~ 0.9 arcsec full-width at half-maximum. After each 120-s integration on the comet, the telescope was moved 3 arcmin south to obtain a sky spectrum, which was subtracted from the original spectrum. The total integration time on the comet was 12 minutes. Signal levels of the individual observations varied by 15%, probably due to errors in guiding the telescope on the comet, although actual brightness variability cannot be ruled out. Errors were estimated after scaling each spectrum to the mean signal level.

The flux-calibrated comet spectrum in a 1.0×5.0 arcsec box ($340 \times 1,700$ km at the comet) is shown in Fig. 1. Division of the comet spectrum by the stellar spectrum did not remove telluric atmospheric absorption features completely, thus the spectrum is

shown uncorrected for the atmosphere. Instead, the atmospheric transmittance was modelled with a radiative transfer program¹⁰. Line fluxes above the atmosphere could then be estimated, taking into account the monochromatic atmospheric transmission at the position of the Doppler-shifted cometary line. The geocentric Doppler shift of the comet at the time was 55.5 km s^{-1} or -0.61 cm^{-1} at $3.0 \mu\text{m}$ and the data are shown without correction for the Doppler shift.

The comet flux (Fig. 1) is mostly the thermal continuum of the coma dust. We identify three emission lines in the ν_3 C–H stretch fundamental of C_2H_2 : the P3, P4 and P5 lines, with rest frequencies given in Table 1. The P3 line is blended with another line. This line and the feature at $3,286.9 \text{ cm}^{-1}$ are OH $v = 1-0$ lines: the Λ -doubled components of the $\text{P}_2(13/2)$ line in the $^2\Pi_{1/2}$ state. Two OH lines in the $v = 2-1$ band are also detected; the Λ -doubled components of the $\text{P}_1(7/2)$ line in the $^2\Pi_{3/2}$ state. The other strong lines in the spectrum are the P8 and P9 lines of the ν_3 C–H stretch fundamental of HCN. The C_2H_2 P4 line has a relatively lower intensity compared to P3 and P5 due to nuclear spin statistics. The factors multiplying the statistical weights have the ratio 1:3 in equilibrium¹¹. The flux in the P4 line is consistent with the equilibrium ratio within the uncertainties.

The dust thermal continuum was clearly extended along the slit with a full-width at half-maximum of ~ 2.2 arcsec. Because of the low contrast, it is difficult to estimate the spatial profiles of the lines. Within the uncertainties, the C_2H_2 and HCN lines had profiles similar to that of the dust, consistent with both being parent molecules released directly from the nucleus. The OH $v = 1-0$ $\text{P}_2(13/2)$ lines were similar, but the $v = 2-1$ $\text{P}_1(7/2)$ lines were visibly extended to the edge of the field of view. The emission lines of C_2H_2 and HCN are the result of resonance fluorescence of sunlight in the $v = 1-0$ vibrational fundamentals. Other possible sources of excitation are believed to be much weaker than infrared solar pumping^{12,13}.

For OH, the situation is more complicated. Both infrared and

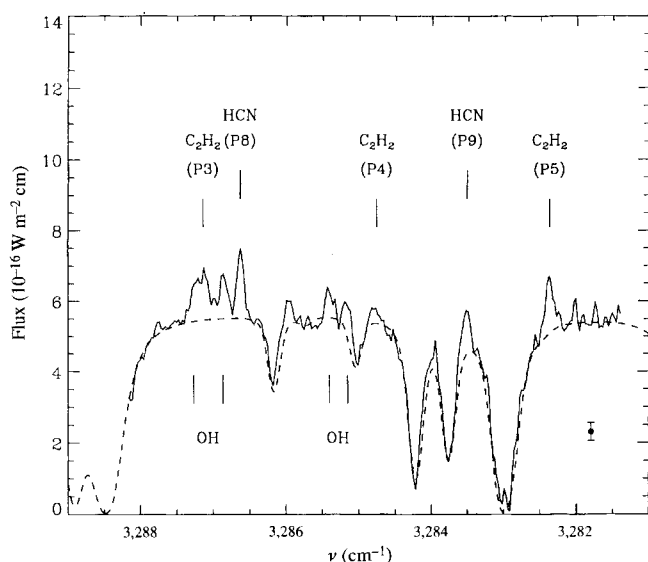


FIG. 1 Spectrum of comet Hyakutake in a 1.0×5.0 arcsec aperture. The spectrum (solid line) is flux-calibrated but not corrected for telluric absorption features. A model of the Earth's atmospheric transmittance at the same spectral resolution (2×10^4) and scaled to match the spectrum is also shown (dashed line). The spectrum was not corrected for the geocentric Doppler shift of the comet ($55.5 \text{ km s}^{-1} = -0.61 \text{ cm}^{-1}$). A typical 1σ error bar is shown at the lower right. The flux calibration was obtained using the star α Persei (spectral type F5Ib, V-band magnitude 1.79), assumed to be a black body of temperature $T = 6,600 \text{ K}$. The loss of starlight in passing through the 1.0 arcsec slit due to seeing was estimated to be 20% for a perfectly centred image and the comet fluxes were corrected for this factor. The uncertainty in the overall flux calibration is estimated to be 15%.

TABLE 1 Observed lines in Hyakutake

Line	Rest frequency (cm^{-1})	F_{line} (above atmosphere) ($10^{-17} \text{ W m}^{-2}$)
C_2H_2 P5	3,282.990	1.5 ± 0.1
C_2H_2 P4	3,285.377	0.5 ± 0.1
C_2H_2 P3	3,287.756	1.1 ± 0.1
HCN P9	3,284.128	1.3 ± 0.1
HCN P8	3,287.248	1.5 ± 0.1
OH $v = 1-0$ $\text{P}_2(13/2)\text{e}$	3,287.478	0.9 ± 0.1
OH $v = 1-0$ $\text{P}_2(13/2)\text{f}$	3,287.877	0.8 ± 0.1
OH $v = 2-1$ $\text{P}_1(7/2)\text{e}$	3,286.010	0.6 ± 0.1
OH $v = 2-1$ $\text{P}_1(7/2)\text{f}$	3,285.764	0.8 ± 0.1

F_{line} refers to a 1.0×2.0 arcsec aperture centred on the nucleus. 1σ uncertainties (in parentheses) do not include overall flux uncertainty of 15%.

ultraviolet pumping can contribute to fluorescence by OH from the $v = 1$ and $v = 2$ levels¹⁴. The observed $v = 1-0$ lines, however, are from a relatively high rotational level, which is not expected to be significantly populated by fluorescence or collisions. A more likely explanation for these lines is prompt emission from OH left in excited vibrational and rotational states after being created in the photodissociation of H_2O (ref. 15). Models of this process in comets are still preliminary, but prompt emission is expected to dominate over fluorescent emission near the nucleus¹⁶. The two processes are predicted to lead to different spatial distributions for the observed lines, with fluorescence following the OH column density and prompt emission the H_2O column density. The CSHELL data are consistent with primarily prompt emission for the $v = 1-0$ $\text{P}_2(13/2)$ lines, but suggest a contribution from fluorescence for the $v = 2-1$ $\text{P}_1(7/2)$ lines.

The emission line fluxes given in Table 1 are for a 1.0×2.0 -arcsec ($340 \times 680 \text{ km}$) box centred on the nucleus. For molecules released directly from the nucleus, the effective volume sampled by this summation along the slit is mostly the inner coma ($< 1,000 \text{ km}$ from the nucleus), where collisions can keep the rotational levels of the ground vibrational states of C_2H_2 and HCN in approximate equilibrium at the local kinetic temperature^{12,13}.

To determine the molecular production rates of C_2H_2 and HCN at the nucleus, we assumed a single temperature for the inner coma and calculated the line fluxes in the solar radiation field¹⁷. The assumed outflow velocity was 0.8 km s^{-1} . The total band emission rate at $r = 1 \text{ AU}$, or g-factor, for the C_2H_2 ν_3 band was revised to $1.5 \times 10^{-4} \text{ s}^{-1}$ based on a more recent band-strength measurement¹⁸.

Gas temperatures in the inner comae of comets are not precisely known. Radio observations of NH_3 , a relatively short-lived parent molecule, in Hyakutake indicate a rotational temperature of 67 K (ref. 19). Infrared observations of H_2O indicate a rotational temperature of 70 K in the inner coma⁴. Rotational temperatures estimated from the C_2H_2 and HCN lines themselves have factor 2 uncertainties but indicate somewhat higher temperatures, between 100 and 150 K . We calculated production rates for coma temperatures of 70 , 100 and 150 K . Quoting the value for 100 K and the range, we find for C_2H_2 the derived production rate is $1.0 \times 10^{27} \text{ molecules s}^{-1}$ (range $(0.8-1.4) \times 10^{27}$). For HCN the derived production rate is $0.8 \times 10^{27} \text{ molecules s}^{-1}$ (range $(0.7-0.9) \times 10^{27}$).

The H_2O production rate at the time of the observations is not precisely known. On 24 March, when the comet was at $r = 1.06 \text{ AU}$, the water production rate was estimated to be $1.7 \times 10^{29} \text{ molecules s}^{-1}$ from infrared observations⁴. Preliminary analysis of Hubble Space Telescope spectra suggests that on 1 April, when the comet was at $r = 0.89 \text{ AU}$, the water production rate was $1.9 \times 10^{29} \text{ molecules s}^{-1}$ (H.A.W., manuscript in preparation). If the water production rate increased as rapidly as r^{-2} from that date, we would expect roughly $2.8 \times 10^{29} \text{ molecules s}^{-1}$, but

preliminary analysis of radio OH observations indicates lower values at the time of the C_2H_2 observations, $\sim 1.5 \times 10^{29}$ molecules s^{-1} on 8 April 14:00 UT (E. Gérard, manuscript in preparation). Thus we adopt a range of $(1.5-3) \times 10^{29}$ molecules s^{-1} for the production rate of H_2O . The relative abundance of C_2H_2 is then 0.3–0.9% with respect to water and 0.2–0.6% for HCN. HCN has been observed at millimetre wavelengths in several comets, including Hyakutake^{20,21}. The derived abundance of HCN from the infrared spectrum is comparable to that inferred from radio observations, 0.16% (ref. 20).

Our detection of C_2H_2 is a new result. Acetylene has long been suspected of being the grandparent of cometary C_2 (ref. 22) (that is, C_2 forms primarily in the photodissociation of an intermediate which is itself formed in the photodissociation of C_2H_2). The C_2 abundance in Hyakutake is estimated to be roughly 0.4% with respect to water²³. There is a difficulty, however, in ascribing all of the C_2 to C_2H_2 in that the photodissociation lifetime for C_2H_2 (primarily to C_2H , though also directly to C_2 with 22% yield), is estimated to be 7.1×10^4 s at $r = 1.0$ AU (ref. 24). But a Haser model analysis of C_2 profiles in comet Halley suggests a grandparent lifetime of only 1.9×10^4 s, assuming a velocity of 0.8 km s^{-1} (ref. 25). Thus acetylene is unlikely to provide all of the observed C_2 ; dust grains and the newly detected C_2H_6 are also likely to contribute.

Current debate about the origins of cometary volatiles centres on the extent to which they represent relatively unaltered interstellar ice grains⁵ or solar nebula material^{6,7}. Hybrid models in which interstellar ices partially evaporate in the protosolar nebula have also been proposed²⁶. Little C_2H_2 is expected from a solar-abundance mixture of gases in thermochemical equilibrium²⁷ yet possibly C_2H_2 could be created by ultraviolet or particle irradiation of CH_4 -containing ices^{5,28}. For comparisons to the interstellar medium, the C_2H_2/CO ratio is also useful. The present results would suggest a C_2H_2/CO ratio in the nucleus of 3.6–12%, assuming a parent CO/H_2O ratio of 5.8% from infrared observations⁴.

C_2H_2 is believed to form in the gaseous interstellar medium by ion–molecule reactions. Measured values of the gas C_2H_2/CO ratio in hot molecular cloud cores (temperature $T = 100$ – $1,000$ K) are of the order of 10^{-3} (ref. 29). Solid C_2H_2 has not yet been detected in the interstellar medium. However, a model for solid phase abundances in molecular clouds at a time when most of the heavy molecules have condensed out gives $C_2H_2/CO = 3\%$ and $C_2H_2/H_2O = 0.3\%$ (ref. 30), in rough agreement with the Hyakutake abundances. It is possible then that the C_2H_2 in Hyakutake formed in the gas phase in a molecular cloud and was brought into the comet nucleus on the surfaces of interstellar grains. For this reason, a careful search for solid C_2H_2 in cold, dense, molecular-cloud regions—for example by the recently commissioned ISO satellite—may prove fruitful. Detection of solid C_2H_2 in a cold cloud in abundances comparable to that of Hyakutake would support the case for comet volatiles being essentially unaltered interstellar ices. □

Received 6 June; accepted 5 September 1996.

1. Krankowsky, D. *et al.* *Nature* **321**, 326–329 (1986).
2. Bockelée-Morvan, D., Colom, P., Crovisier, J., Despois, D. & Paubert, G. *Nature* **350**, 318–320 (1991).
3. Mumma, M. J., Weissman, P. R. & Stern, S. A. in *Protostars and Planets III* (eds Levy, E. H. & Lunine, J. I.) 1177–1252 (Univ. Arizona Press, Tucson, 1993).
4. Mumma, M. J. *et al.* *Science* **272**, 1310–1314 (1996).
5. Greenberg, J. M. in *Comets* (ed. Wilkening, L. L.) 131–163 (Univ. Arizona Press, Tucson, 1982).
6. Fegley, B. & Prinn, R. G. in *The Formation and Evolution of Planetary Systems* (eds Weaver, H. A. & Danly, L.) 171–211 (Cambridge Univ. Press, Cambridge, 1989).
7. Lunine, J. I. in *The Formation and Evolution of Planetary Systems* (eds Weaver, H. A. & Danly, L.) 213–242 (Cambridge Univ. Press, Cambridge, 1989).
8. Tokunaga, A. T., Toomey, D. W., Carr, J., Hall, D. N. B. & Epps, H. W. *Proc. SPIE* **1235**, 131–143 (1990).
9. Greene, T. P., Tokunaga, A. T., Toomey, D. W. & Carr, J. S. *Proc. SPIE* **1946**, 313–324 (1993).
10. Meadows, V. S. & Crisp, D. J. *Geophys. Res.* **101**, 4595–4622 (1996).
11. Herzberg, G. *Molecular Spectra and Molecular Structure Vol II, Infrared and Raman Spectra of Polyatomic Molecules* (Van Nostrand Reinhold, New York, 1945).
12. Crovisier, J. & Encrenaz, Th. *Astron. Astrophys.* **126**, 170–182 (1983).
13. Weaver, H. A. & Mumma, M. J. *Astrophys. J.* **276**, 782–797 (1984).
14. Schleicher, D. G. & A'Hearn, M. F. *Astrophys. J.* **331**, 1058–1077 (1988).
15. Crovisier, J. *Astron. Astrophys.* **213**, 459–464 (1989).

16. Bockelée-Morvan, D. & Crovisier, J. *Astron. Astrophys.* **216**, 278–283 (1989).
17. Crovisier, J. *Astron. Astrophys. Suppl. Ser.* **68**, 223–258 (1987).
18. Vander Auwera, J., Hurtmans, D., Carier, M. & Herman, M. J. *Mol. Spectrosc.* **157**, 337–357 (1993).
19. Wootton, A. *et al.* *Abstr. Int. Conf. on Asteroids, Comets, Meteors 39* (Observatoire de Paris, Meudon, 1996).
20. Matthews, H. E. *et al.* *IAU Circ. No.* 6353 (1996).
21. Womack, M., Festou, M. C. & Stern, S. A. *IAU Circ. No.* 6345 (1996).
22. Jackson, W. M. J. *Photochem.* **5**, 107–118 (1976).
23. Schleicher, D. *IAU Circ. No.* 6344 (1996).
24. Huebner, W. F., Keady, J. J. & Lyon, S. P. *Astrophys. Space Sci.* **195**, 1–294 (1992).
25. O'Dell, C. R., Robinson, R. R., Krishna Swamy, K. S., McCarthy, P. J. & Spinrad, H. *Astrophys. J.* **334**, 476–488 (1988).
26. Yamamoto, T. in *Comets in the Post-Halley Era* (eds Newburn, R. L. Jr, Neugebauer, M. & Rahe, J.) 361–376 (Kluwer, Dordrecht, 1991).
27. Prinn, R. G. & Fegley, B. *Astrophys. J.* **249**, 308–317 (1981).
28. Strazzulla, G. & Johnson, R. E. in *Comets in the Post-Halley Era* (eds Newburn, R. L. Jr, Neugebauer, M. & Rahe, J.) 243–275 (Kluwer, Dordrecht, 1991).
29. Carr, J. S., Evans, N. J., Lacy, J. H. & Zhou, S. *Astrophys. J.* **450**, 667–690 (1995).
30. Hasegawa, T. I. & Herbst, E. *Mon. Not. R. Astron. Soc.* **263**, 589–606 (1993).

ACKNOWLEDGEMENTS. T.Y.B. and A.T.T. are visiting Astronomers of the NASA Infrared Telescope Facility, which is operated by the University of Hawaii under contract to NASA. We thank the staff of the IRTF and R. Joseph for helping to make these observations successful on short notice, and M. J. Mumma and M. A. DiSanti for critical readings of the manuscript.

CORRESPONDENCE should be addressed to T.Y.B. (e-mail: tyb@keajpl.nasa.gov).

Photoelectrochromic windows and displays

Clemens Bechinger*, Suzanne Ferrere, Arie Zaban, Julian Sprague & Brian A. Gregg

National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, Colorado 80401-3393, USA

PHOTOCHROMIC materials^{1,2} change colour on absorption of light, whereas electrochromic materials^{3,4} change colour in response to an electrically induced change in oxidation state. Both classes of materials are being investigated for potential applications in displays, imaging devices and 'smart' windows^{5–8,15,16}. Here we describe an alternative route to such applications, in which an electrochromic film and a photovoltaic film form the two electrodes of an electrochemical cell. The resulting structure exhibits photochromism, but unlike conventional photochromic films, the light-absorption process (in the photovoltaic film) is separate from the coloration process (in the electrochromic film): both may therefore be optimized individually. Moreover, as the coloration process in our cells requires an external electrical current between the two electrodes, the optical state of the cell—transparent, absorbing or, in the case of non-uniform illumination, patterned—can be stored when the circuit is open, or changed when the electrodes are connected.

The light-absorbing function in the photoelectrochromic (PEC) cell is performed by a dye-sensitized semiconductor electrode that produces a photovoltage sufficient to colour the electrochromic film deposited on the counter electrode. Dye-sensitization has a long history of use in colour photography and is currently being studied as a potential means of solar energy conversion^{9–11}. Some of the best-known electrochromic materials are inorganic oxides such as WO_3 and MoO_3 . We illustrate the principle of the photoelectrochromic cells with a ruthenium polypyridine-sensitized nanocrystalline TiO_2 electrode, similar to those introduced by Grätzel *et al.* for use in solar cells^{9,10}, and a WO_3 electrochromic film on the counter electrode.

The dye, $Ru(II)L_2L'$, where L is 2,2'-bipyridine-4,4'-dicarboxylate and L' is 4,4'-dimethyl-2,2'-bipyridine, was adsorbed onto 4- μm -thick nanocrystalline TiO_2 films^{9,10} from dilute (8 μM) ethanol solution containing 30 mM chenodeoxycholic acid. The dye coverage was kept low to make the cells almost transparent in the 'off' state; the coadsorbate, chenodeoxycholic acid, was used to

* Permanent address: Universität Konstanz, Postfach 5560 M675, D-78434 Konstanz, Germany.

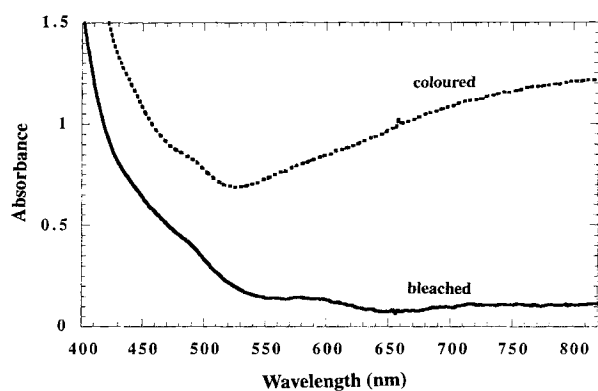
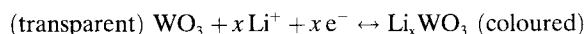


FIG. 1 Absorption spectra of a PEC cell before (bleached) and after (coloured) illumination at an intensity of $\sim 75 \text{ mW cm}^{-2}$ for 1 min. The absorption maximum of the dye occurs at 480 nm but is partially obscured by scattering from the TiO_2 electrode. The tungsten bronze (Li_xWO_3) that is responsible for most of the absorbance in the coloured state has a broad absorption maximum around 1,000 nm.

increase the photocurrent efficiency¹². The counter electrode was prepared by thermal evaporation of a 500-nm-thick layer of WO_3 onto an indium-tin oxide coated glass substrate. The two electrodes were bonded to each other around their periphery with 30- μm -thick strips of Surllyn 1601. The gap between the electrodes was filled, through small holes drilled in the WO_3 electrode, with propylene carbonate containing 0.1 M LiI and 0.01 M 4-*t* butylpyridine. The fill holes were then sealed with epoxy cement. Cells were fabricated ranging in size from 1 to 25 cm^2 .

Light absorption by the sensitizing dye leads to electron injection into the TiO_2 film, with a quantum yield at short circuit of $\sim 70\%$. This is followed by re-reduction of the oxidized dye via electron transfer from I^- in solution. At open circuit, the dye-sensitization process produces a photovoltage of 0.6–0.9 V but does not result in coloration of the cell. When the electrodes are short-circuited, however, the electrons move from the TiO_2 through the external circuit to the WO_3 film where, in the presence of cations small enough to intercalate into the WO_3 lattice, a bluish-coloured tungsten bronze is formed. When the cation is Li^+ , for example, the reaction leading to coloration of the WO_3 film is¹³



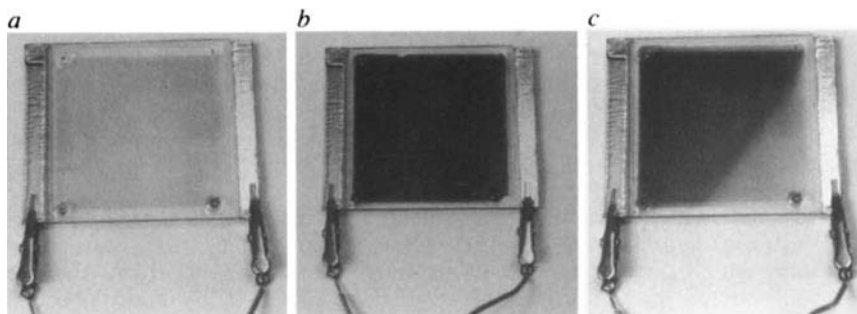
The electrochromic process is highly reversible and lifetimes of over 10^7 cycles have been achieved in display applications³. The coloration process is equivalent to charging a battery; the photovoltage generated by the dye-sensitized electrode is used to drive Li^+ ions and electrons into the WO_3 film. When the light is blocked after the WO_3 electrode is charged, a voltage equal to the open circuit photovoltage, but of opposite polarity, appears across the PEC cell. If the electrodes are then short-circuited, this voltage causes the cell to discharge spontaneously back to its bleached

state. If the electrodes are opened before the incident light is blocked, the cell remains coloured because the electrons cannot escape from the reduced WO_3 film. Figure 1 shows the absorption spectra of a PEC cell before and after illumination for 1 min at $\sim 75 \text{ mW cm}^{-2}$ intensity (comparable to noon daylight at sea level, 45° latitude), illustrating the bleached and the coloured states, respectively.

Photographs of a 25- cm^2 cell in bleached, coloured and inhomogeneously coloured states are shown in Fig. 2. When the electrolyte layer is thin ($\sim 25 \mu\text{m}$ in our examples) compared to the size of the cell, spatially resolved coloration occurs when just part of the cell is illuminated. Lateral ion mobility in the electrolyte limits the spatial resolution of the device; however, once the Li^+ ions are intercalated into the WO_3 film, their lateral mobility becomes so small that an image can be stored for many hours. Figure 2c shows an example where only part of the device was illuminated. The unexposed part remained transparent whereas the illuminated part became coloured. The pattern was still visible after storage of the device with open-circuited terminals for more than 24 h. The rate of colouring and bleaching of a cell on exposure to chopped white light of $\sim 75 \text{ mW cm}^{-2}$ intensity was probed by measuring the transmission of a weak laser beam at 788 nm wavelength (Fig. 3). On illumination, the transmission decreased to its saturation level after ~ 100 s. Recovery to the bleached state after the light was blocked was somewhat slower.

The PEC devices appear to have advantages over earlier designs of 'smart' windows^{3,5,6}. Use of smart windows can significantly decrease the air-conditioning costs in commercial buildings⁶. However, existing smart windows require an external power source to change the transmittance, making the retro-fitting of existing buildings difficult and expensive. A self-powered smart window, which uses the energy of the incident sunlight to modulate its own transmissivity, could be easily installed in existing buildings without requiring additional electrical connections. One recently proposed approach is a photovoltaic-electrochromic window consisting of a thin-film amorphous silicon carbide photovoltaic cell connected in tandem with an electrochromic window^{6,14}. It consists of eight sequentially deposited layers that must be extremely thin in order for the device to be semitransparent. Unfortunately, electrical short-circuits are a serious problem and it has proved difficult to demonstrate a large-area device. Dye-sensitized electrodes have several distinct advantages over conventional photovoltaic cells, or other semiconductor electrodes, for large-area window or display applications. The PEC cells described here are remarkably simple devices that are easily assembled without electrical short-circuits. The light-absorbing dye layer can be made optically thin by either reducing the thickness of the TiO_2 layer or by decreasing the concentration of the adsorbed dye. On the other hand, decreasing the optical thickness of inorganic semiconductor films often leads to electrical short-circuits. Inorganic semiconductors absorb all wavelengths that correspond to energies greater than their band-gap, and thus can be transparent only if the bandgap lies in the ultraviolet range. In contrast, sensitizing dyes of any colour, even infrared-absorbing dyes, can be employed for dye-sensitized

FIG. 2 Photographs of a PEC cell with a 25 cm^2 active area. *a*, The bleached state. The cell is light yellow owing to the sensitizing dye and some scattering from the TiO_2 electrode. *b*, The uniformly coloured state achieved by illumination at an intensity of $\sim 75 \text{ mW cm}^{-2}$ for 1 min. The colour is dark blue. *c*, The inhomogeneously coloured state achieved by illumination with one corner of the cell masked. The states shown in *b* and *c* remain approximately unchanged for at least 24 h if the cell is kept at open circuit.



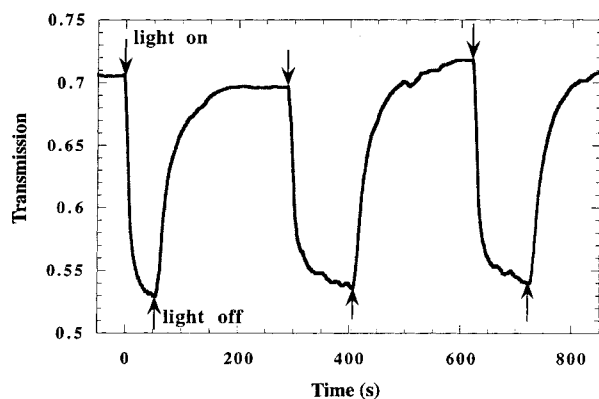


FIG. 3 Change in transmission at 788 nm wavelength of a short-circuited PEC cell exposed to chopped white light.

electrodes. In the latter case, the device would be transparent to visible light in the 'off' state but could be darkened by sunlight or addressed with an infrared diode laser. Perhaps most importantly,

the dye-sensitized electrodes are inexpensive, robust and can be easily scaled up to give large-area substrates. □

Received 5 August; accepted 9 September 1996.

1. Dürr, H. & Bouas-Laurent, H. *Photochromism: Molecules and Systems* (Elsevier, Amsterdam, 1990).
2. Yao, J. N., Hashimoto, K. & Fujishima, A. *Nature* **355**, 624–626 (1992).
3. Deb, S. K. *Sol. Energy Mater. Sol. Cells* **25**, 327–338 (1992).
4. Green, M., Smith, W. C. & Weiner, J. A. *Thin Solid Films* **38**, 89–100 (1976).
5. Granqvist, C. G. *Solid State Mater. Sci.* **16**, 291–308 (1990).
6. Benson, D. K. & Branz, H. M. *Sol. Energy Mater. Sol. Cells* **39**, 203–211 (1995).
7. Nagasu, M. & Koshida, N. *Appl. Phys. Lett.* **57**, 1324–1325 (1990).
8. Deb, S. K. *Appl. Opt. Suppl.* **3**, 192–195 (1969).
9. O'Regan, B. & Grätzel, M. *Nature* **353**, 737–740 (1991).
10. Nazeeruddin, M. K. et al. *J. Am. Chem. Soc.* **115**, 6382–6390 (1993).
11. Parkinson, B. A. & Spiller, M. T. *Electrochim. Acta* **37**, 943–948 (1992).
12. Kay, A. & Grätzel, M. *J. Phys. Chem.* **97**, 6272–6277 (1993).
13. Faughnan, B. W., Crandall, R. S. & Heyman, P. M. *RCA Rev.* **36**, 177–197 (1975).
14. Bechinger, C. et al. *J. Appl. Phys.* **80**, 1226–1232 (1996).
15. Bonhôte, P. et al. *Chem. Commun.* 1163–1164 (1996).
16. DeBerry, D. W. & Viehbeck, A. J. *Electrochem. Soc.* **130**, 249–251 (1983).

ACKNOWLEDGEMENTS. We thank S.-Y. Huang for technical assistance and helpful discussions; this work was supported by the US Department of Energy. We thank the following for support: the Office of Energy Research, Division of Basic Energy Sciences, Advanced Energy Projects Division (C.B., S.F.); the Office of Energy Efficiency and Renewable Energy, Office of Utility Technologies, Photovoltaics Division (A.Z.); and the Office of Energy Research, Division of Basic Energy Sciences, Chemical Sciences Division (J.S., B.G.).

CORRESPONDENCE and requests for materials should be addressed to B.A.G. (e-mail: bgregg@nrel.nrel.gov).

The three-dimensional structure of an upper ocean vortex in the tropical Pacific Ocean

Pierre J. Flament*, Sean C. Kennan*, Robert A. Knox†, Pearn P. Niiler† & Robert L. Bernstein‡

* School of Ocean and Earth Science and Technology, University of Hawaii, Honolulu, 1000 Pope Road, Hawaii 96822, USA

† Scripps Institution of Oceanography, University of California San Diego, San Diego, La Jolla, California 92093-0230, USA

‡ SeaSpace, Inc., 9240 Trade place suite 100, San Diego, California 92126, USA

In the equatorial Pacific Ocean, easterly trade winds and the Earth's rotation combine to drive surface currents away from the Equator, thereby causing cold nutrient-rich subsurface water to upwell. The front¹ that forms between this upwelled water and warmer waters north of the Equator is sometimes visible as a spectacular 'line in the sea'² between 2° and 6° N. Westward-propagating cusp-shaped disturbances observed along this front³ have been attributed to the effect of dynamical instabilities in the system of zonal equatorial currents^{4–11} but the connection between these phenomena remains unclear. Here we report extensive measurements from shipboard sensors, satellite and drifting buoys which reveal the three-dimensional structure of an anticyclonic eddy (or vortex) 500 km in diameter and centred at 4° N. We suggest that cusp-shaped disturbances of the front are caused by trains of large-amplitude vortices, which are driven by instability of the mean zonal shear. We show that these vortices not only play an important role in the meridional transport of heat, salt and momentum, but are also associated with regions of intense horizontal convergence along the front, where dramatic concentrations of marine life are observed.

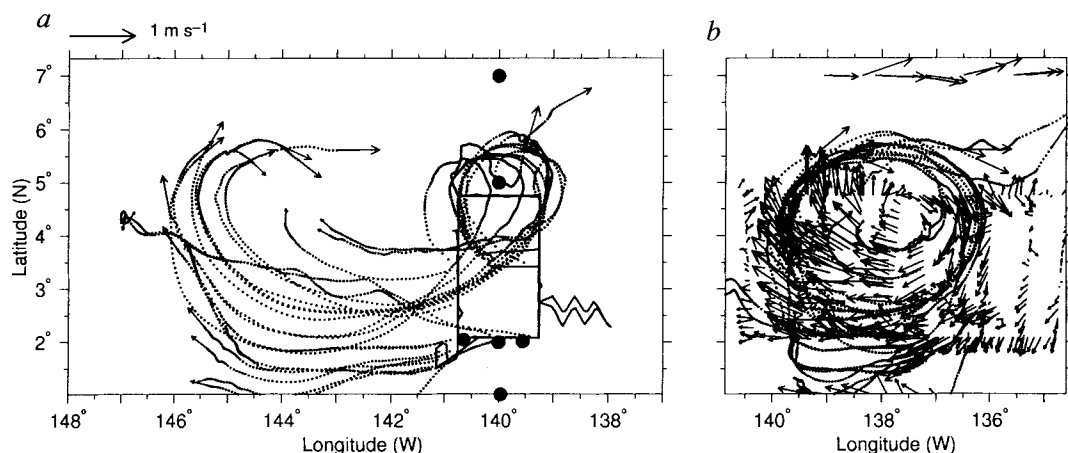
The sheared zonal equatorial currents system (ECS), consisting of the westward south equatorial current (SEC) at 2° S to 3° N, the eastward equatorial undercurrent (EUC) and the eastward north equatorial counter-current (NECC) at 5° N–10° N, is subject to westward propagating perturbations at periods of 15 to 30 days,

first observed as meanders of the Atlantic equatorial currents⁴. They were subsequently identified in satellite images of sea surface temperature (SST)³, drifting buoy trajectories⁵, sea level and dynamic topography^{32,33}, shipboard acoustic doppler current profiler (ADCP) sections⁶, equatorial moored current meters time series^{7,8}, and numerical model simulations^{9–11}. The amplitude of these perturbations is largest when the ECS strengthens seasonally in boreal summer and autumn. They result presumably from dynamical instability of the ECS, although conversion of mean to eddy energy appears to occur at several latitudes, varying through the annual cycle^{12,13}.

Repeated hydrographic surveys centred at 140° W, 3° 25' N were conducted in November 1990, with the objective of understanding the off-equatorial dynamics of these perturbations. The thermohaline fields were sampled with conductivity–temperature–depth (CTD) stations to 300 m depth spaced at 40 km, and near the fronts with CTD sensors attached to a towed platform (SeaSoar), undulating over 4 km (ref. 14). A 150-kHz ship-mounted ADCP yielded profiles of velocity and backscatter intensity to 300 m depth; absolute velocities were computed using global-positioning system (GPS) navigation and the ship's gyrocompass. Forty satellite-tracked (ARGOS) buoys¹⁵ drogued at 15 m depth were also deployed. Meteorological sensors were mounted on a bow tower to estimate the air–sea fluxes of heat, moisture and momentum by bulk formulae^{16,17}. Additional measurements of wind, current and temperatures were obtained from moorings deployed by others^{18,19}.

The drifting buoys followed typical cycloidal trajectories, often observed in this region²⁰, with a meridional amplitude of ~400 km, a zonal wavelength of ~600 km, and an orbital period of 20 days (Fig. 1a). Cycloids result from the superposition of a rotation and a translation; the translation speed inferred from the trajectories is 30 cm s⁻¹ westward. Viewed in coordinates moving at this speed, velocity measurements from ADCP and moorings line up remarkably well with the buoy trajectories (Fig. 1b), despite their wide spatio-temporal separation in fixed coordinates. This indicates that the flow was approximately steady in the translating coordinates. The velocity, temperature and salinity fields were therefore mapped in these coordinates, with a median interpolation of 70 km. The translation speed is robust: minimization of the interpolation error, fits to individual trajectories, and maximization of the correlation between longitude and meridional velocity,

FIG. 1 a, Trajectories of the buoys between 17 November and 5 December 1990, dotted every 6 h. The last points are marked by velocity arrows. The track of the ship is shown by a solid line, and the positions of the moorings by filled circles. b, Trajectories of the buoys in a frame of reference translating westward at 30 cm s^{-1} , with velocity vectors from ADCP and moorings superimposed (the longitude scale corresponds to the position of the vortex on 16 November 1990 at 00:00 UTC).



ADCP and mooring data included, yield the same speed, with a standard error of 1 cm s^{-1} .

The interpolated surface velocity (Fig. 2a) reveals a 500-km diameter anticyclonic vortex drifting along $4^{\circ} 24' \text{ N}$, straddled by the $\sim 0.60 \text{ m s}^{-1}$ eastward NECC north of $6^{\circ} 30' \text{ N}$, and by the $\sim 0.75 \text{ m s}^{-1}$ westward SEC south of $2^{\circ} 30' \text{ N}$. The zonal section through the centre (Fig. 3a) is asymmetric: the northward flow near 140° W peaked at 60 cm s^{-1} near the surface, whereas the southward flow near 137° W peaked at 75 m depth. The vortex was confined to the upper layer; below the 150-m deep thermocline, velocities were less than 10 cm s^{-1} . The density field was consistent with geostrophic balance, reflected by a deepening of the pycnocline in the core.

Remotely sensed SST (Fig. 2a) and zonal sections of temperature and salinity (Fig. 3b, c) show northward transport of equatorial water (relatively cold and saline from the upwelling), and southward transport of a shallow layer of warmer and fresher water (influenced by precipitation in the inter-tropical convergence zone, ITCZ). Cold and warm waters were separated by a sharp front, advected meridionally by the circulation of the vortex. The cusps generally seen in SST images³ appear therefore to be the expression of such off-equatorial vortices, deforming the north equatorial front.

The spatial structure of the fields in the translating coordinates is equivalent to a time series along a meridian. As about one period was sampled, the contribution of this vortex to meridional fluxes can be estimated by zonal averaging. At the surface, the heat flux (Fig. 4c) was southward and the salt flux (Fig. 4d) was northward everywhere, down the mean meridional gradients. The flux of zonal momentum (Fig. 4e) was negative south of

5° N , that is, there was a northward flux of westward SEC momentum; it diverged south of, and converged north of $3^{\circ} 40' \text{ N}$ —about the latitude of the average position of the front. The large meridional momentum flux between 2° N and 5° N , acting to reduce the shear between the SEC and the NECC, suggests that the vortex was the finite amplitude result of an off-equatorial instability of the mean shear, in agreement with numerical model predictions^{10,11,21}.

During the ship surveys, the three moorings deployed along the Equator¹⁹ recorded a meridional velocity oscillation of $\sim 70 \text{ cm s}^{-1}$ amplitude and 18 d period. Cross-correlations of meridional velocity between adjacent equatorial moorings yielded an unambiguous westward propagation speed of 78 cm s^{-1} , in contrast to the 30 cm s^{-1} drift speed inferred for the off-equatorial vortex. This suggests that the vortex was distinct, and possibly decoupled, from the equatorial oscillations, known to be excited by shear within the SEC and between the SEC and UEC^{12,13,19}.

Convergence of the large-scale wind-driven surface flow (the Ekman flow)²² favours the formation of SST fronts. The curl of the zonally averaged wind stress (Fig. 4f) is negative between $2^{\circ} 30' \text{ N}$ and $4^{\circ} 30' \text{ N}$, as southeasterly trade winds decrease towards the ITCZ. The Ekman flow is thus convergent, sharpening temperature and salinity gradients, and eventually leading to the formation of fronts. Similar Ekman fronts are observed in the North Pacific²³ and North Atlantic^{24,25} subtropical gyres.

Horizontal velocity gradients (that is, vorticity, divergence) offer further insight into the frontogenesis dynamics. The potential vorticity Q of a column of water, related to its angular momentum, is the ratio of the sum of planetary vorticity f (the vertical component of the Earth's rotation) and relative vorticity ζ

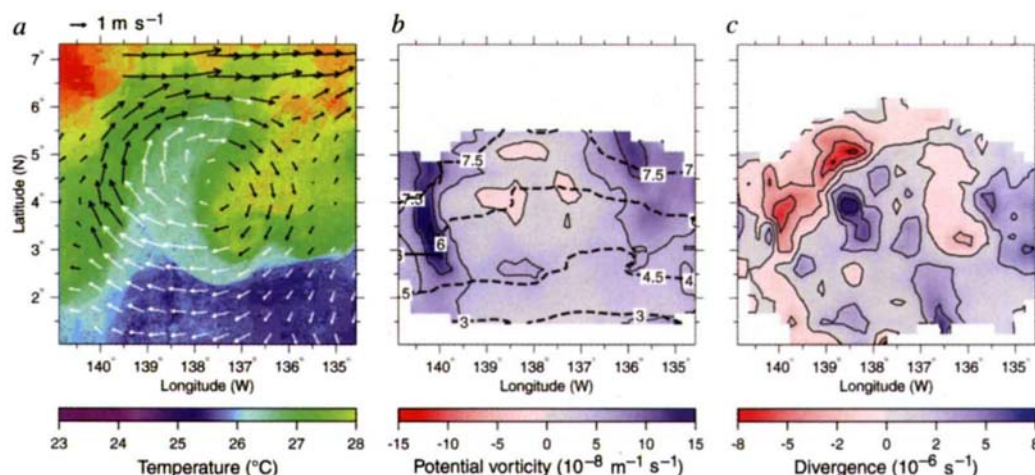


FIG. 2 Fields in the moving frame of reference. a, Interpolated surface velocity; the background shows sea-surface temperature from a satellite image on 16 November at 23:20 UTC; b, total potential vorticity for the upper layer (the planetary component is contoured with dashed lines); c, horizontal divergence at the surface.

(the rotation of the vortex with respect to the Earth), to the thickness H of the column: $Q = (f + \zeta)H^{-1}$. Potential vorticity of the upper layer (Fig. 2b) was conserved along streamlines, despite a twofold variation of planetary vorticity between 3° N and 6° N; relative vorticity was an important ingredient in this conservation. The near-zero vorticity in the core suggests that the growth of the vortex may have been limited by inertial instabilities.

Vertical velocity, downwelling or upwelling, is directly related to horizontal velocity divergence (Fig. 2c). A long band of convergence, stretching from 140° W 3° 30' N to 138° W 6° N, coincided with the leading edge of the front; it extended down the thermocline, implying vertical motions reaching 32 m d⁻¹. A similar pattern of convergence was predicted by numerical models¹¹. The convergence band is consistent with the conservation of potential vorticity: as water columns move rapidly northward, their planetary vorticity f increases; to conserve potential vorticity Q , their thickness H must also increase (that is, the flow converges), further sharpening the Ekman front.

Convergence increases gradients of other advected properties such as zooplankton concentration, which can be estimated from the acoustic backscatter intensity of the ADCP²⁶. A zonal section of backscatter anomaly (Fig. 3d), constructed using only daytime data to avoid bias due to the diurnal migration of zooplankton, shows an astounding increase of more than 20 dB above the background, aligned with the northward converging flow at 139° 30' W. In the southward flow, the scattering layer was less intense and at the depth of the thermocline, suggesting that scatterers are subducted downwards while advecting around the vortex.

These new observations provide a synoptic synthesis of the causal links between large-scale atmospheric and oceanic flows, dynamical instabilities, finite amplitude vortices, frontogenetic processes, and eventual concentration of biological organisms. Intensification of the equatorial front in the presence of a vortex²⁷, increased density of phytoplankton^{2,28} and zooplankton²⁹ at the surface, and large accumulation of phytodetritus³⁰ at the sea bottom, were observed again during the 1992 Equatorial Pacific Joint Global Ocean Flux Experiment (JGOFS-EQPAC)³¹, consistent with our findings from data collected in 1990.

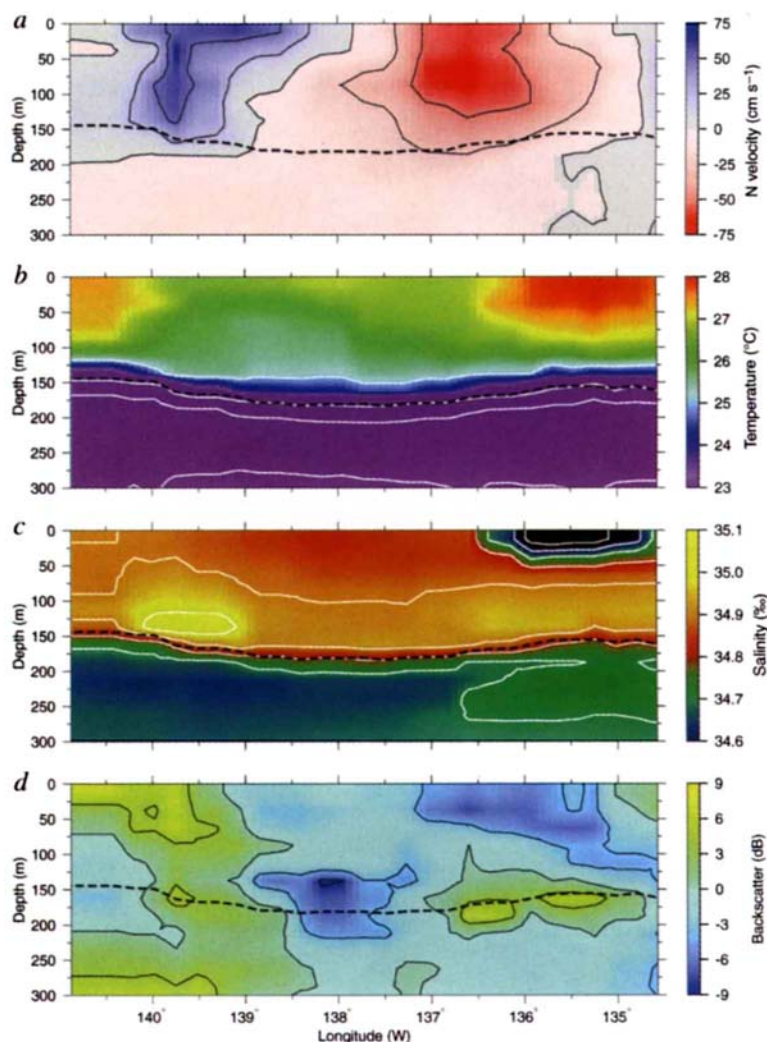


FIG. 3 Zonal section through the vortex along 4° 24' N: a, meridional (northward) velocity component; b, temperature; c, salinity; d, daytime acoustic backscatter at 150 kHz, normalized by the cruise-average profile. The 1,024.5 kg m⁻³ isopycnal, conventionally chosen to identify the depth of the thermocline, is shown by a dashed line.

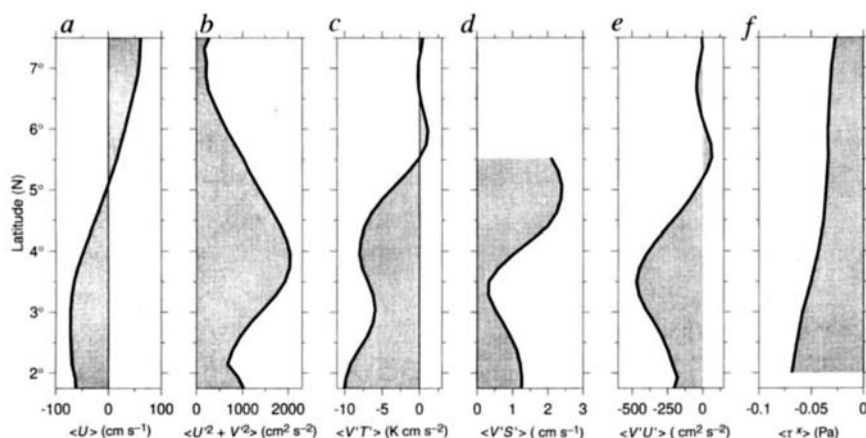


FIG. 4 Zonally averaged quantities: a, b, mean zonal component of velocity (U); a, b, mean zonal component of eddy kinetic energy ($\langle U^2 + V^2 \rangle$); c, d, e, meridional fluxes of heat ($\langle V'T' \rangle$); c, salt ($\langle V'S' \rangle$); d, and zonal component of momentum ($\langle U'V' \rangle$); e; f, mean zonal wind stress (τ).

Westward propagating perturbations of the ECS have been referred to as 'equatorial long waves' (Rossby gravity waves causing equatorial meanders of the EUC/SEC⁸), 'Legeckis waves' (cusps in the north equatorial front, seen in satellite images³), and often, somewhat interchangeably, as 'tropical instability waves'. Our results, integrating data from a rich range of sensors, indicate that away from the Equator, they consist of finite amplitude vortices, drifting along the shear layer between the SEC and the NECC. Although when sampled from a fixed point, they appear as periodic oscillations of the fields, just as linear waves would, they are clearly trains of fully developed large-amplitude nonlinear eddies.

Moreover, within the limitations of the short sampling time, the observation of drastically different propagation speeds along the Equator and along 4.5° N, suggests that the equatorial and sub-equatorial perturbations may at some times appear as distinct phenomena, contrary to general assumption. Whether this results from different forcing mechanisms, which may be phase-locked at some times of the year and decoupled at others (the vortex observed here was the last of the 1990 season), or from separate manifestations of the same instability, raises new questions concerning the dynamics of the equatorial current system in the Pacific. □

Received 15 December 1995; accepted 30 August 1996.

1. Knauss, J. A. *Tellus* **IX**, 234–237 (1957).
2. Yoder, J. A., Ackleson, S. G., Barber, R. T. & Balch, W. M. *Nature* **371**, 689–692 (1994).
3. Legeckis, R. *Science* **197**, 1179–1181 (1977).
4. Düing, W. et al. *Nature* **257**, 280–284 (1975).
5. Hansen, D. V. & Paul, C. A. *J. Geophys. Res.* **89**, 10431–10440 (1984).
6. Wilson, D. & Leetmaa, A. *J. Phys. Oceanogr.* **18**, 1641–1657 (1988).
7. Philander, S. G. *EOS* **66**, 154 (1985).
8. Halpern, D., Knox, R. A. & Luther, D. S. *J. Phys. Oceanogr.* **18**, 1514–1534 (1988).
9. Philander, S. G. *J. Geophys. Res.* **81**, 3725–3734 (1976).
10. Cox, M. D. *J. Phys. Oceanogr.* **10**, 1168–1186 (1980).
11. Philander, S. G. H., Hurlin, W. J. & Pacanowski, R. C. *J. Geophys. Res.* **91**, 14207–14211 (1986).
12. Luther, D. S. & Johnson, E. S. *J. Phys. Oceanogr.* **7**, 913–944 (1990).
13. Johnson, E. S. & Luther, D. S. *J. Geophys. Res.* **99**, 7689–7705 (1994).
14. Pollard, R. T. *Nature* **323**, 433–435 (1986).
15. Niiler, P. P., Davis, R. E. & White, H. J. *Deep-Sea Res.* **34**, 1867–1882 (1987).
16. Smith, S. D. *J. Geophys. Res.* **93**, 15467–15472 (1988).
17. Liu, W. T., Katsaros, K. B. & Businger, J. A. *J. Atmos. Sci.* **36**, 1722–1735 (1979).
18. McPhaden, M. J. *J. Geophys. Res.* **101**, 6337–6359 (1996).
19. Qiao, L. & Weisberg, R. H. *J. Geophys. Res.* **100**, 8677–8693 (1995).

20. Chew, F. & Bushnell, M. H. *J. Phys. Oceanogr.* **20**, 1124–1133 (1990).
21. Donohue, K. A., thesis, Univ. Rhode Island (1996).
22. Price, J. F., Weller, R. A. & Schudlich, R. R. *Science* **238**, 1534–1538 (1987).
23. Niiler, P. P. & Reynolds, R. W. *J. Phys. Oceanogr.* **14**, 217–230 (1984).
24. Pollard, R. T. & Regier, L. *Nature* **348**, 227–229 (1990).
25. Weller, R. *Nature* **348**, 199–200 (1990).
26. Flagg, C. N. & Smith, S. L. *Deep-Sea Res.* **36**, 455–474 (1989).
27. Johnson, E. S. *Deep-Sea Res.* (in press).
28. Foley, D. G. et al. *Geophys. Res. Lett.* (submitted).
29. Landry, M. R., Kirshtein, J. & Constantinou, J. *Deep-Sea Res.* (submitted).
30. Smith, C. R. et al. *Deep-Sea Res.* (submitted).
31. Sawyer, M. & Flament, P. *Deep-Sea Res.* **42**, (1995); cover.
32. Pengaud, C. *J. Geophys. Res.* **95**, 7239–7248 (1990).
33. Miller, L., Watts, D. R. & Wimbush, M. *J. Phys. Oceanogr.* **15**, 1759–1770 (1985).

ACKNOWLEDGEMENTS. We thank the captain and crew of the RV Moana Wave; K. Constantine, J. Firing, H. Ramm, M. Sawyer and C. Trefois for assisting with data collection and processing; and M. Landry and D. Luther for providing comments on the manuscript. This work was supported by the US National Science Foundation and by the State of Hawaii.

CORRESPONDENCE should be addressed to P.F. (e-mail: flament@soest.hawaii.edu).

Effects of social organization on gene flow in the fire ant *Solenopsis invicta*

D. DeWayne Shoemaker* & Kenneth G. Ross†

* Department of Biology, University of Rochester, Rochester, New York 14627, USA

† Department of Entomology, University of Georgia, Athens, Georgia 30602-2603, USA

A CONTROVERSIAL model of speciation proposes that the development of alternative social organizations within populations of group-living animals may drive the inception of reproductive isolation^{1–3}. The alternative social behaviours, which are selectively favoured in some social or ecological contexts, may be correlated with distinctive reproductive traits such that significant barriers to interbreeding emerge between coexisting social variants. Evidence for this mode of speciation is almost non-existent^{3–8}, but it provides one of the most compelling mechanisms for sympatric speciation^{3,8} and could conceivably explain many species origins. Here we examine variation in mitochondrial DNA and two unique nuclear genes to demonstrate that gene flow between sympatric social forms of the fire ant *Solenopsis invicta* is restricted to only one of four possible routes. The loss of the other routes results from incompatibilities in the social systems of the two forms, demonstrating the potential for social selection to generate significant barriers to gene flow and to initiate reproductive isolation.

The fire ant *Solenopsis invicta* includes two distinctive social forms that differ in the number of egg-laying queens per nest; nests of the monogyne (M) form contain only a single queen, whereas those of the polygyne (P) form may contain 200 or more queens⁸. The two forms occur sympatrically in many parts of the native range in South America⁹ as well as in the introduced range

in the United States⁸, where polygyny is thought to have arisen repeatedly from monogyny under conditions of high local population density⁸. The social forms differ in several features of their breeding biology, which are associated with colony queen number and are expected to influence gene flow between the forms. The differences include the site of mating, extent of dispersal and mode of colony founding^{8,10}. We use markers of the mitochondrial and nuclear genomes to infer how these differences affect gene flow between the social forms where they occur in sympatry in the introduced range.

Fire ants were sampled from an area of northern Georgia, USA, where the two forms co-occur (Fig. 1). Restriction-fragment length polymorphism analysis of a maternally inherited 4-kilobase mitochondrial DNA (mtDNA) fragment led to the detection of four composite haplotypes (designated A, B, C and D). An analysis of the haplotype variance¹¹ reveals that only a relatively small amount (<14%) occurs among sampling sites within each form, with such among-site differentiation significant only in the P form (Table 1). This confinement of among-site differentiation to the P form accords with the suggestion that queens of the P form are less vagile than queens of the M form (ref. 8 and K.G.R. and D.D.S., manuscript submitted). The majority of haplotype variation (>50%) occurs between the social forms, with this interform differentiation being statistically significant using either of two genetic distance metrics (Table 1). Thus, queen-mediated gene flow is greatly restricted between the social forms of *S. invicta*.

A site-by-site comparison of the haplotype frequencies (Table 2) reveals the magnitude of queen-mediated interform gene flow occurring via different routes. One common haplotype in the P form (C) is lacking in the M form, even at the Eastville site immediately downwind of the main area of polygyny, a pattern consistent with the complete absence of female migration from the P form to the M form. Two haplotypes with moderate pooled frequencies in the M form (B and D) are virtually absent in the P form. Only two of 230 P queens, both from the Walton Academy site at the windward boundary of the main P population with the M form, possess either of these haplotypes (Table 2). Their complete absence from the other Georgia P sites, as well as

TABLE 1 Hierarchical analysis of molecular variance for mtDNA haplotypes of *S. invicta*

Variance component	Equidistant metric			Euclidean metric		
	% of total variance	Φ -statistics	<i>P</i>	% of total variance	Φ -statistics	<i>P</i>
All samples						
Between social forms	51.4	$\Phi_{FT} = 0.514$	0.020	55.9	$\Phi_{FT} = 0.559$	0.030
Among sites						
Within forms	13.6	$\Phi_{SF} = 0.280$	<0.001	13.4	$\Phi_{SF} = 0.303$	<0.001
All sites	(65.0)	$\Phi_{ST} = 0.650$	<0.001	(69.3)	$\Phi_{ST} = 0.693$	<0.001
Within sites	35.0			30.7		
Monogyne (M) samples						
Among sites	0	$\Phi_{SF} = 0$	0.444	0.5	$\Phi_{SF} = 0.005$	0.294
Within sites	100			99.5		
Polygyne (P) samples						
Among sites	36.1	$\Phi_{SF} = 0.361$	<0.001	35.8	$\Phi_{SF} = 0.358$	<0.001
Within sites	63.9			64.2		

The equidistant metric assumes that all four haplotypes are equally divergent from one another, whereas the euclidean metric is a distance measure that equals the number of restriction site differences between haplotypes. The total haplotypic variance is partitioned among three levels in the upper part of the table: between forms, among sites and within sites. The proportion of variance among all sites (in parentheses) is the sum of the variance between forms and the variance among sites within forms. The statistic Φ_{FT} measures the extent of differentiation between forms, Φ_{SF} measures differentiation among sites within each form, and Φ_{ST} measures differentiation among all sites. The probabilities (*P*) that the estimates of the Φ -statistics differ from zero (no differentiation) were determined by permutation analysis using 1,000 randomly permuted data matrices²¹.

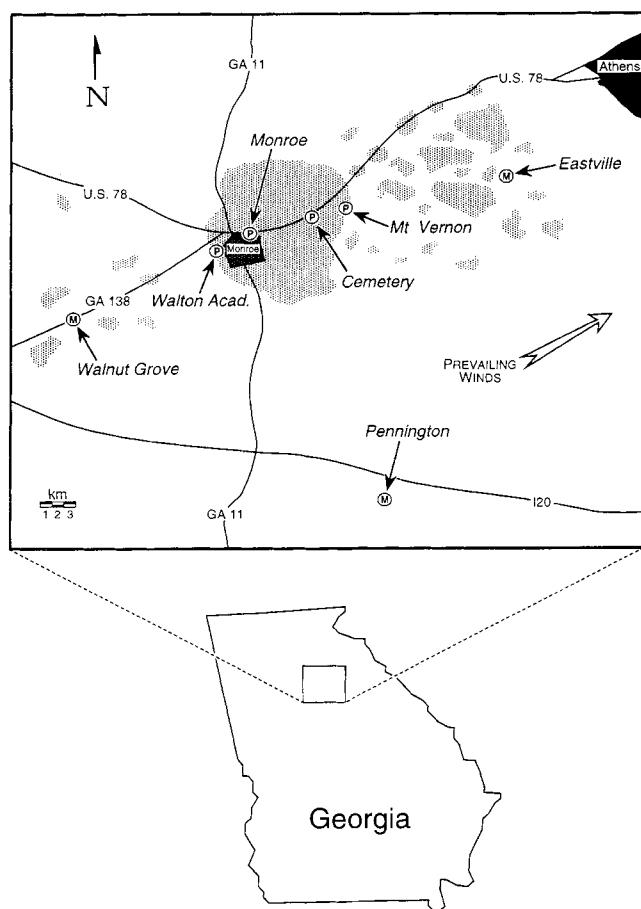


FIG. 1 Locations of seven sampling sites for monogyne (M) and polygyne (P) *S. invicta* in northern Georgia, USA. Areas in which only P nests occur are indicated by stippling; other areas contain mostly or only M nests. All 149 nests sampled at the three M sites were assigned to the monogyne form on the basis of diagnostic nestmate genotype arrays at five polymorphic allozyme loci^{19,24,28}. All 230 nests sampled at the four P sites were assigned to the polygyne form on the basis of the discovery of more than one egg-laying (wingless) queen in each. The mean direction of the prevailing winds during the main mating flights is shown¹⁵.

from a P site in Mississippi (K.G.R. and D.D.S., manuscript submitted), indicates that haplotype B occurs at the Walton Academy site because of a small amount of mtDNA leakage from the M to the P form at their interface. The pronounced mtDNA differentiation observed between the forms at our Georgia study locality is significant because nests at many of the sites are within a single dispersal distance of nests of the alternative form¹².

The importance of interform gene flow that occurs by means of dispersing P males mating with M queens was examined by studying allele frequencies at two nuclear loci under different selection regimes in the two forms. The alleles *Pgm-3^b* and *Gp-9^b* are maintained at moderate to high frequency in the P form because of strong directional selection or overdominance; these alleles are much less common (*Pgm-3^b*) or absent (*Gp-9^b*) in the M form, where such selection pressures are absent (refs 13, 14 and K.G.R., manuscript submitted). Frequencies of *Pgm-3^b* in the inferred mates of M queens are virtually identical to the frequencies in the queens and in wild-caught males sampled from the same M sites, but differ strongly from the frequencies in queens and wild-caught males from P nests (Fig. 2). Moreover, *Gp-9^b* was inferred to be completely absent in the mates of M queens, even though 50% of males from P nests are expected to carry it (K.G.R., manuscript submitted). Thus, allele frequencies in the mates of M queens at these two nuclear loci are those expected if such males derive exclusively from M nests. This conclusion again is particularly significant for the M site directly downwind of the main area of polygyny (Eastville).

These data demonstrate that three of four potential routes of gene flow do not lead to significant movement of genes between the social forms of *S. invicta*. That the final route is important in this respect is indicated by previous results showing that P queens commonly mate with and use the sperm of M males. For instance, *Pgm-3^b* and *Gp-9^b* frequencies inferred for the mates of P queens are those expected if M males predominate among these mates (ref. 10 and K.G.R., manuscript submitted), and the estimated proportion of such interform matings drops off with increasing distance from M nests¹⁵. Substantial nuclear gene flow occurring via this route also explains the fact that P populations throughout the introduced range in the US closely resemble adjacent M populations in their allele frequencies at a diversity of polymorphic nuclear loci^{10,16}.

The two social forms of *S. invicta* are thus linked by an unusual

TABLE 2 Frequencies of four mtDNA haplotypes in the monogyne and polygyne forms of *S. invicta*

	Haplotypes			
	A	B	C	D
Monogyne (M) form				
Walnut Grove	0.837	0.163	0	0
N = 43	(0.720–0.930)	(0.070–0.280)		
Pennington	0.894	0.053	0	0.053
N = 38	(0.868–1.0)	(0–0.132)		(0–0.132)
Eastville	0.867	0.118	0	0.015
N = 68	(0.779–0.941)	(0.044–0.191)		(0–0.044)
All Sites	0.866	0.114	0	0.020
N = 149	(0.805–0.919)	(0.060–0.168)		(0–0.047)
Polygyne (P) form				
Walton Academy	0.159	0.032	0.809	0
N = 63	(0.079–0.254)	(0–0.079)	(0.746–0.921)	
Monroe	0.167	0	0.833	0
N = 54	(0.074–0.259)		(0.741–0.926)	
Cemetery	0.057	0	0.943	0
N = 53	(0–0.132)		(0.868–1.0)	
Mt Vernon	0.700	0	0.300	0
N = 60	(0.583–0.800)		(0.200–0.417)	
All Sites	0.280	0.007	0.713	0
N = 230	(0.222–0.335)	(0–0.022)	(0.652–0.770)	

The 95% confidence intervals about the estimates (in parentheses) were obtained by drawing 1,000 bootstrap replicates from the original data sets and eliminating the 25 extreme high and 25 extreme low frequencies derived from these replicates. *N* is the number of nests (= individuals) sampled.

pattern of gene flow in which one sex mediates gene flow in only one direction. This pattern can be explained on the basis of the contrasting breeding systems and reproductive phenotypes distinguishing the two forms^{8,10,15}. M queens disperse widely during mating flights and found new colonies independently (without workers), relying on extensive nutrient reserves acquired during adult maturation to accomplish this task¹⁷. In contrast, P queens mate in their natal nest, where they attempt to become egg layers^{18,19}, or they participate in mating flights and seek adoption into other P nests^{18,20}. In either case, they initiate egg laying in an existing colony, so that large nutrient reserves are not necessary. Indeed, excessive reserves are disadvantageous to queens attempting to become egg layers in P nests, because workers of this form are intolerant of queens with 'high-reserves' phenotypes¹⁴. Thus, queens are not expected to be important agents of interform gene flow because M queens possess an inappropriate phenotype for securing adoption into P nests and the colonies they found independently invariably become monogyne^{14,21}, and because P queens generally do not have sufficient reserves to start monogyne colonies independently and they are not accepted into established M colonies^{22,23}. Males of the P form also are not expected to mediate significant interform gene flow (by mating with M queens during mating flights), because most P males (>90%) are sterile²⁴ and fertile P males may specialize in within-nest mating⁸. The remaining route of gene flow, M males mating with P queens, is expected to be important, as has been found, because operational sex ratios are highly female biased in the P form but slightly male biased in the M form²⁵. Thus, M males that disperse into P populations should achieve high mating success because of the continual availability of receptive P queens.

Despite considerable gene flow by a single route, the two social forms of *S. invicta* may be vulnerable to eventual cessation of all genetic exchange, as might occur for instance if mating flights were completely abandoned by P queens in favour of within-nest mating (loss of the mating flight coupled with adoption of within-nest mating seems to be a common evolutionary trend in polygyne ants^{8,26}). From this perspective, the loss of most routes of gene flow between the fire ant social forms may illustrate important steps in the development of complete reproductive isolation between conspecific populations that have diverged in social organization^{1,28}. These ants thus provide the most direct evidence

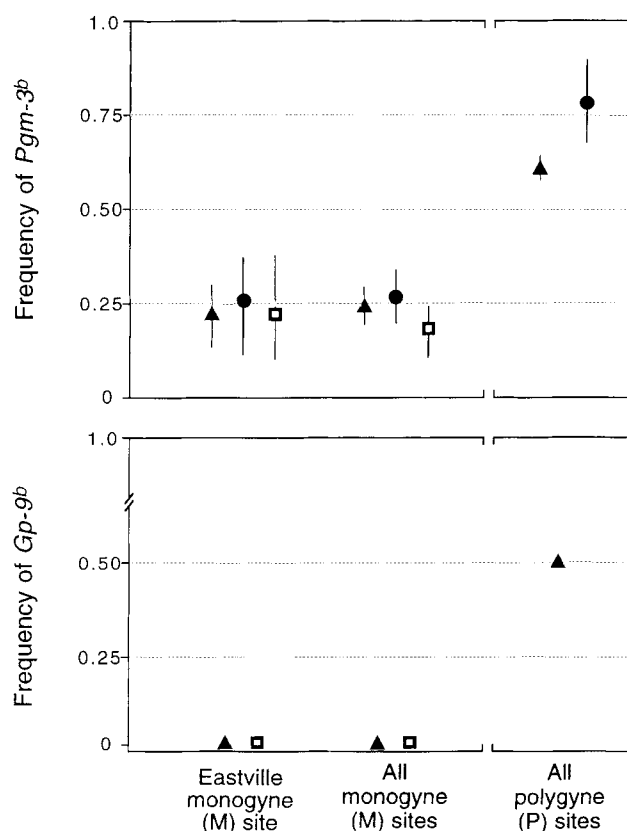


FIG. 2 Frequencies of the alleles *Pgm-3^b* and *Gp-9^b* in egg-laying queens (filled triangles) of the two social forms and in the inferred mates of monogyne (M) queens (open squares) of *S. invicta*. Also shown for comparison are *Pgm-3^b* frequencies in wild-caught males of the two forms (filled circles). *Gp-9* is not expressed in males (K.G.R., manuscript submitted), so data for this gene are not available from wild-caught males. However, because virtually all egg-laying polygyne (P) queens are *Gp-9^b* heterozygotes (K.G.R., manuscript submitted), 50% of the fertile males produced in P nests are expected to bear allele *Gp-9^b* in the hemizygous condition. Bars represent 95% confidence intervals about the frequency estimates derived from 1,000 bootstrap replicates (see Table 2).

to date for the feasibility of sympatric speciation driven by social selection. □

Methods

Mitochondrial DNA analyses. A 4-kb portion of the mtDNA (including the A + T-rich noncoding region) was amplified by polymerase chain reaction (PCR) from a single queen per nest and digested with the restriction enzymes *Bam*HI, *Dpn*II, *Eco*RV, *Eco*RI, *Hae*III, *Hha*I, *Hind*III, *Hinf*I, *Kpn*I, *Msp*I, *Rsa*I, *Taq*I and *Xba*I (K.G.R. and D.D.S., manuscript submitted). Digestion products were separated electrophoretically in 1.5% agarose gels, stained with ethidium bromide, and visualized using ultraviolet light. To confirm that the desired region of the mtDNA molecule was amplified, one end of the PCR product was sequenced and aligned with a previously published sequence from the honey bee²⁷. Maternal inheritance of the 4-kb fragment was demonstrated by examining the variable digestion products from the enzyme *Taq*I in nestmates from ten monogynous colonies. The presence or absence of restriction sites inferred using complete and partial digestion procedures defined the composite haplotypes.

Estimation of *Pgm-3^b* and *Gp-9^b* frequencies. The genotypes of M queens and their single mates were inferred by inspecting genotype arrays for female and male offspring from single M nests. Such reconstruction of parental genotypes is possible because a queen's sons arise from her unfertilized eggs and her daughters arise from eggs fertilized by the sperm of a single haploid male^{24,28}. For P queens and wild-caught males, allele frequencies were estimated using a resampling procedure when more than one individual was collected per nest¹⁵. Wild-caught males from P nests were confirmed to be fertile haploids on the basis of their size and their banding patterns at four allozyme loci¹³. Methods for scoring genotypes at *Pgm-3* and *Gp-9* and evidence for Mendelian inheritance of the products of these genes are presented elsewhere (refs 13, 28 and K.G.R., manuscript submitted). Sample sizes for estimating *Pgm-3^b* and *Gp-9^b* frequencies are as follows. Eastville M site: *Pgm-3^b*, egg-laying queens and their male mates from 41 nests, 149 wild-caught males from 19 nests; *Gp-9^b*, egg-laying queens and their male mates from 67 nests. All M sites: *Pgm-3^b*, egg-laying queens and their male mates from 118 nests, 497 wild-caught males from 62 nests; *Gp-9^b*, egg-laying queens and their male mates from 149 nests. All P sites: *Pgm-3^b*, egg-laying queens from

427 nests, 140 wild-caught males from 28 nests; *Gp-9^b*, egg-laying queens from 427 nests.

Received 16 July; accepted 27 August 1996.

- West-Eberhard, M. J. *Q. Rev. Biol.* **58**, 155–183 (1983).
- West-Eberhard, M. J. *Proc. Natl Acad. Sci. USA* **83**, 1388–1392 (1986).
- Bourke, A. F. G. & Franks, N. R. *Biol. J. Linn. Soc.* **43**, 157–178 (1991).
- Crozier, R. H. *Annu. Rev. Entomol.* **22**, 263–288 (1977).
- Buschinger, A. *Trends Ecol. Evol.* **1**, 155–160 (1986).
- Ward, P. S. in *The Genetics of Social Evolution* (eds Breed, M. D. & Page, R. E.) 123–148 (Westview, Boulder, 1989).
- Buschinger, A. *Z. Zool. Syst. Evol. Forsch.* **28**, 241–260 (1990).
- Ross, K. G. & Keller, L. *Annu. Rev. Ecol. Syst.* **26**, 631–656 (1995).
- Ross, K. G., Vargo, E. L. & Keller, L. *Evolution* (in the press).
- Ross, K. G. & Shoemaker, D. D. *Evolution* **47**, 1595–1605 (1993).
- Excoffier, L., Smouse, P. E. & Quattro, J. M. *Genetics* **131**, 470–491 (1992).
- Markin, G. P., Dillier, J. H., Hill, S. O., Blum, M. S. & Hermann, H. R. *J. Georgia Entomol. Soc.* **6**, 145–156 (1971).
- Ross, K. G. *Nature* **355**, 347–349 (1992).
- Keller, L. & Ross, K. G. *Science* **260**, 1107–1110 (1993).
- Ross, K. G. & Keller, L. *Am. Nat.* **146**, 325–348 (1995).
- Ross, K. G., Vargo, E. L. & Fletcher, D. J. C. *Evolution* **41**, 979–990 (1987).
- Markin, G. P., Collins, H. L. & Dillier, J. H. *Ann. Entomol. Soc. Am.* **65**, 1053–1058 (1972).
- Porter, S. D. *J. Entomol. Sci.* **26**, 474–478 (1991).
- Ross, K. G., Vargo, E. L. & Keller, L. *Proc. Natl Acad. Sci. USA* **93**, 3021–3025 (1996).
- Glancey, B. M. & Lofgren, C. S. *Florida Entomol.* **71**, 581–587 (1988).
- Tschinkel, W. R. & Howard, D. F. *Behav. Ecol. Sociobiol.* **12**, 103–113 (1983).
- Fletcher, D. J. C. & Blum, M. S. *Science* **219**, 312–314 (1983).
- Porter, S. D., Van Eimeren, B. & Gilbert, L. E. *Ann. Entomol. Soc. Am.* **81**, 913–918 (1988).
- Ross, K. G., Vargo, E. L., Keller, L. & Trager, J. C. *Genetics* **135**, 843–854 (1993).
- Vargo, E. L. *J. Evol. Biol.* (in the press).
- Bourke, A. F. G. & Franks, N. R. *Social Evolution in Ants* (Princeton Univ. Press, Princeton, NJ, 1995).
- Crozier, R. H. & Crozier, Y. C. *Genetics* **133**, 97–117 (1993).
- Shoemaker, D. D., Costa, J. T. & Ross, K. G. *Heredity* **69**, 573–582 (1992).

ACKNOWLEDGEMENTS. We thank J. Avise, C. DeHeer, J. Evans, M. Goodisman, L. Keller, M. A. Moran and E. Vargo for comments on the manuscript. This research was funded by grants from the National Geographic Society and the Georgia Agricultural Experiment Stations, University of Georgia. D.D.S. was supported by a grant from the US Department of Agriculture to J. Jaenike.

CORRESPONDENCE and requests for materials should be addressed to K.G.R. (e-mail: ross@bscr.uga.edu).

Word meanings can be accessed but not reported during the attentional blink

Steven J. Luck*, Edward K. Vogel* & Kimron L. Shapiro†

* Department of Psychology, University of Iowa, Iowa City, Iowa 52242-1407, USA

† School of Psychology, University of Wales, Bangor, Gwynedd LL57 2DG, UK

AFTER the detection of a target item in a rapid stream of visual stimuli, there is a period of 400–600 ms during which subsequent targets are missed. This impairment has been labelled the ‘attentional blink’¹. It has been suggested that, unlike an eye blink, the additional blink does not reflect a suppression of perceptual processing, but instead reflects a loss of information at a postperceptual stage, such as visual short-term memory^{2–4}. Here we provide electrophysiological evidence that words presented during the attentional blink period are analysed to the point of meaning extraction, even though these extracted meanings cannot be reported 1–2 s later. This shows that the attentional blink does indeed reflect a loss of information at a postperceptual stage of processing, and provides a demonstration of the modularity of human brain function.

To test whether the attentional blink reflects a postperceptual impairment in processing, we recorded event-related potentials (ERPs) from normal young adults and measured the ‘N400’ peak, (a negative peak at 400 ms poststimulus) which reflects the degree of mismatch between a word and a previously established semantic context^{5–7}. For example, a large N400 would be elicited by the

last word of the sentence ‘The man wore blue trousers and a green bucket’, but not by the last word of the sentence ‘The man wore blue trousers and a green shirt’. Because a word must be identified before it can be compared with a semantic context, the presence of an N400 peak for a mismatching word indicates that this word has been identified. Thus the presence of a normal N400 in our study would provide strong evidence that words presented during the attentional blink are fully identified, even though subjects cannot accurately report them.

As shown in Fig. 1, each trial in this experiment began with the presentation of a ‘context word’ that created the semantic context for that trial. A stream of 20 seven-character strings of letters or numbers was then presented at a rapid rate of one string every 83 ms. Most were distractor items consisting of randomly selected consonants, drawn in blue. Either the seventh or the tenth string served as the first target on a given trial, and this string consisted of a randomly selected digit, repeated seven times (to create a seven-character string). The second target consisted of a word of 3–7 characters, drawn in red. We will refer to this second target as the ‘probe’ word. At the end of the trial, a question mark appeared, signalling the subject to make two button-press responses. These responses indicated whether the first target was an odd digit or an even digit and whether the probe word was semantically related or unrelated to the context word for that trial. The context word and probe word were either closely related (for example, jam and jelly) or clearly unrelated (for example, shoe and mouse). A control condition was also included in which subjects ignored the first target.

The probe word was always the first, third or seventh string after the first target (denoted below as lag 1, lag 3 and lag 7). Previous studies have shown that the attentional blink is typically strongest at lags 2–3 and usually ends by lags 6–8, with little or no impairment in accuracy at lag 1 (refs 1–3). As shown in Fig. 2a, we found the same pattern of results. Specifically, accuracy for the

probe word was 87–90% correct at lags 1 and 7, but dropped to 66% correct at lag 3. Accuracy for identifying the relatedness of the probe word was higher at all lags in the control condition than in the experimental condition and showed no decline at lag 3, resulting in giving a statistically significant main effect of condition ($P < 0.001$) and a significant interaction between lag and condition ($P < 0.001$).

Our main question was whether the probe-elicited N400 peak, like probe discrimination accuracy, would also be reduced at lag 3 in the experimental condition. The analysis of the N400 was complicated by the high stimulation rate, which caused the ERP elicited by the probe word to be overlapped by the ERPs elicited by the previous and subsequent stimuli. To avoid this problem, we constructed 'difference waves' in which the ERPs elicited by probe words that were related to the initial context word were subtracted from the ERPs elicited by unrelated probe words. Because these two trial types were identical except for the relationship between the probe word and the context word, only the potentials that were influenced by this relationship (the N400) remained in the difference wave.

As shown in Fig. 3, the difference waves consisted primarily of a negative-going potential that had the typical characteristics of the N400 component^{4,8}: it peaked at roughly 400 ms, was largest at the central and parietal midline electrodes, and was slightly larger over the right hemisphere than the left. The N400 was significantly smaller in the experimental condition than in the control condition ($P = 0.01$), which reflects the lower overall accuracy in this condition and demonstrates the sensitivity of the N400. As shown in Fig. 2b, however, there was no significant effect of lag on N400 amplitude in either the experimental condition or the control condition, despite the large drop in accuracy at lag 3 in the experimental condition ($P > 0.35$ for both the lag main effect and the lag $\times$ condition interaction).

Thus, although the N400 seems to be a sensitive index of semantic-mismatch detection, the attentional blink was not accompanied by a reduction in the N400 elicited by a semantic

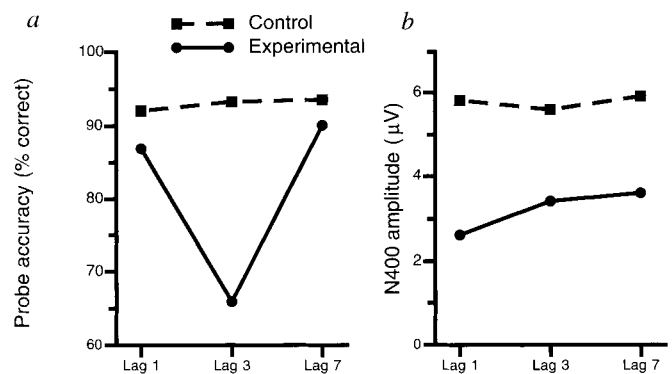


FIG. 2 a, Probe-discrimination accuracy as a function of lag (lag of 1-, 3- or 7-strings) for the experimental and control conditions. These values reflect only the trials on which the first target was correctly discriminated (first-target accuracy was 96% correct overall, with no effect of lag). b, Mean N400 amplitude as a function of lag for probe words in the experimental and control conditions, measured from the unrelated – related difference waves and averaged across electrode sites. N400 amplitude was computed as the mean amplitude of electrical activity between 300 and 500 ms poststimulus, relative to a 200-ms prestimulus baseline, at the F3, Fz, F4, C3, Cz, C4, P3, Pz and P4 electrode sites.

mismatch. This is especially remarkable given that we observed a substantial decrease in the ability of our subjects to report this semantic mismatch at the end of the trial, 1–2 s after the peak of the N400. This dissociation between N400 amplitude and discrimination accuracy implies that many of the errors at lag 3 in the experimental condition occurred despite correct semantic analysis, which should in turn have led to the presence of an N400 on these error trials. This prediction was tested in a separate analysis of the ERPs elicited on trials with incorrect responses for the probe word. In the experimental condition, the N400 on these error trials was found to be significantly larger at lag 3 than at lags 1 or 7 ($P < 0.05$). Together with the lack of a suppression at lag 3 on correct trials, these results provide strong evidence that the decline in probe discrimination accuracy during the attentional blink was not caused by a suppression in perceptual processing, but was instead a result of a postperceptual loss of information.

Our findings have implications beyond the simple question of whether the attentional blink operates at an early or late stage. First, the finding of no perceptual suppression during the attentional blink period contrasts with the results of spatial-attention experiments, in which the N400 peak and the sensory-evoked P1 and N1 peaks (the initial positive and negative ERP peaks) are suppressed for stimuli presented at ignored locations compared with attended locations^{9–12}. This suggests that the attentional mechanisms that underlie the selection of perceptual information from a spatial array are different from the mechanisms that underlie the selection of more highly processed conceptual information over a period of time. This independence of perceptual and postperceptual processes may be important for tasks that involve selection at both low and high levels, such as reading, because it may allow the faster low-level perceptual processes to operate asynchronously from the slower high-level processes¹³. These findings also contribute to our understanding of the general architecture of the human cognitive system, by indicating that the meaning of a word can be extracted and compared with other semantic information without reaching a stage at which the results of this comparison can be retained for even 1–2 s. This may further indicate that substantial semantic processing can occur in the absence of awareness, although it is difficult to determine whether the probe words were identified without reaching awareness or if they momentarily reached awareness and were then rapidly forgotten. □

Stimulus type	Time (ms)	Related trial	Unrelated trial
Context word	1,000	RAZOR	WHEEL
Blank	1,000		
Distractor	83	PNVCSZP	KDSWPVZ
Distractor	83	GRSDPKN	VNMCPKL
Distractor	83	BVCPLMS	FDPMCNV
Distractor	83	DSPWTFR	VPMTDZM
Distractor	83	RLDJHKG	HJDLGFP
Distractor	83	SPLDJMF	DFPLJKH
First target	83	3333333	4444444
Distractor	83	WDPTBNF	GHJDMVT
Distractor	83	SCDPVBF	HDVCBNM
Probe	83		
Distractor	83	FDLNLKB	NMCVPHJ
Distractor	83	DLJJCNW	DCVPBJM
Distractor	83	WPSCDSN	PCNBVLK
Distractor	83	DPWVCPB	NPMTVDK
Distractor	83	CBNDPNJ	BRTFPMF
Distractor	83	RTPMVBC	JLSDCDK
Distractor	83	TWISCLMN	LKSDVCP
Distractor	83	LJVBCMH	DKKHNVF
Distractor	83	RMVCPKL	WKLDMZP
Distractor	83	DPNMNVZ	CPNHVGB
Blank	1,000		
Response cue	2,000	?	?
Blank	2,000		

FIG. 1 Example stimulus sequences for trials on which the probe word was either related or unrelated to the context word. The probe word was presented in red and all other items were drawn in blue. Also note that the probe word was flanked by Xs, when necessary, to create a total of seven characters in the string.

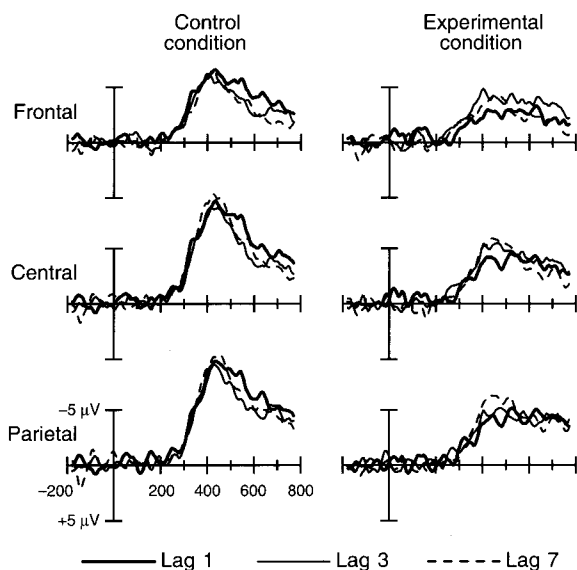


FIG. 3 ERP difference waves at frontal, central and parietal electrode sites along the midline (Fz, Cz and Pz), averaged across the 14 subjects. These waveforms were produced from averages that included only the trials on which the first target was correctly discriminated, but were not sorted according to the accuracy of the response to the probe word. The waveforms were low-pass filtered by convolving them with a gaussian impulse-response function with a standard deviation of 10 ms and a 50% amplitude cutoff of 20 Hz. Time zero represents the onset of the probe word. Note that, by convention, negative is plotted upwards.

Methods

Each string subtended $4.9^\circ \times 0.8^\circ$ of visual angle. The first target was equally likely to be an odd or even digit, and the probe word was equally likely to be related or unrelated to the context word. Each subject received six 60-trial blocks in the experimental condition (in which both targets required a response) and six 60-trial blocks in the control condition (in which only the probe word required a response) in counterbalanced order. ERPs were recorded from 14 right-handed, neurologically normal native English speakers using our standard recording procedures¹¹ and electrodes located at 15 standard electrode sites (International 10/20 system) spanning the scalp. Analysis of variance was used for all statistical tests.

Received 8 July; accepted 3 September 1996.

1. Raymond, J. E., Shapiro, K. L. & Arnell, K. M. *J. Exp. Psychol. Hum. Percept. Perform.* **18**, 849–860 (1992).
2. Shapiro, K. L., Raymond, J. E. & Arnell, K. M. *J. Exp. Psychol. Hum. Percept. Perform.* **20**, 357–371 (1994).
3. Chun, M. M. & Potter, M. C. *J. Exp. Psychol. Hum. Percept. Perform.* **21**, 109–127 (1995).
4. Shapiro, K., Driver, J., Ward, R. & Sorensen, R. E. *Psychol. Sci.* (in the press).
5. Kutas, M. & Hillyard, S. A. *Science* **270**, 203–205 (1980).
6. Besson, M., Kutas, M. & Van Petten, C. *J. Cog. Neurosci.* **4**, 132–149 (1992).
7. Osterhout, L. & Holcomb, P. J. in *Electrophysiology of Mind* (eds Rugg, M. D. & Coles, M. G. H.) 171–215 (Oxford Univ. Press, New York, 1995).
8. Kutas, M., Van Petten, C. & Besson, M. *Electroenceph. Clin. Neurophysiol.* **69**, 218–233 (1988).
9. McCarthy, G. & Nobre, A. C. *Electroenceph. Clin. Neurophysiol.* **88**, 210–219 (1993).
10. Mangun, G. R., Hillyard, S. A. & Luck, S. J. in *Attention and Performance XIV* (eds Meyer, D. & Kornblum, S.) 219–243 (MIT Press, Cambridge, Massachusetts, 1993).
11. Luck, S. J. et al. *J. Exp. Psychol. Hum. Percept. Perform.* **20**, 887–904 (1994).
12. Eimer, M. *Percept. Psychophys.* **55**, 667–675 (1994).
13. Potter, M. C. *J. Exp. Psychol. Hum. Learn. Mem.* **2**, 509–522 (1976).

ACKNOWLEDGEMENTS. This work was supported by a grant from the McDonnell-Pew Program in Cognitive Neuroscience.

CORRESPONDENCE and requests for materials should be addressed to S.J.L. (e-mail: steven-luck@uiowa.edu).

Role of posterior parietal cortex in the recalibration of visually guided reaching

Dottie M. Clower*, John M. Hoffman*†, John R. Votaw†, Tracy L. Faber†, Roger P. Woods‡ & Garrett E. Alexander*

* Department of Neurology and † Department of Radiology, Emory University School of Medicine, 1639 Pierce Drive, P.O. Drawer V, Atlanta, Georgia 30322, USA

‡ Division of Brain Mapping, UCLA School of Medicine, Los Angeles, California 90095, USA

VISUALLY guided reaching requires complex neural transformations to link visual and proprioceptive inputs with appropriate motor outputs^{1,2}. Despite the complexity of these transformations, hand-eye coordination in humans is remarkably flexible, as demonstrated by the ease with which reaching can be adapted to distortions in visual feedback. If subjects attempt to reach to visual targets while wearing displacing prisms, they initially misreach in the direction of visual displacement. Given feedback about their reaching errors, however, they quickly adapt to the visual distortion. This is shown by the gradual resumption of accurate reaching while the prisms remain in place, and by the immediate onset of reaching errors in the opposite direction after the prisms have been removed³. Despite an abundance of psychophysical data on adaptation to prisms, the functional localization of this form of sensorimotor adaptation is uncertain. Here we use positron emission tomography (PET) to localize changes in regional cerebral blood flow (rCBF) in subjects who performed a prism-adaptation task as well as a task that controlled for the sensory, motor and cognitive conditions of the adaptation experi-

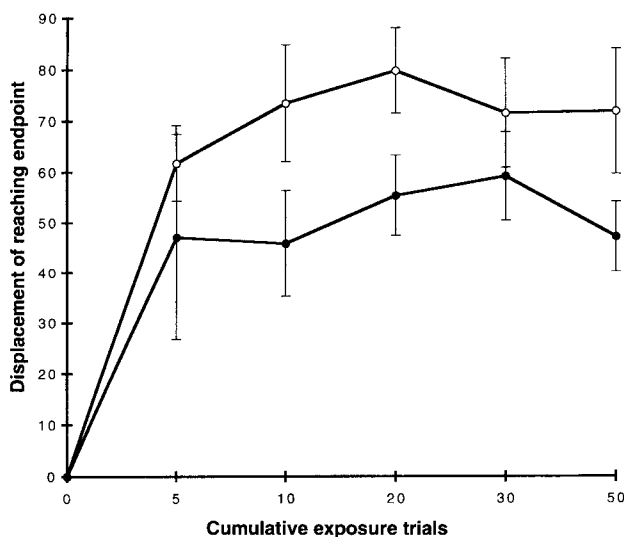
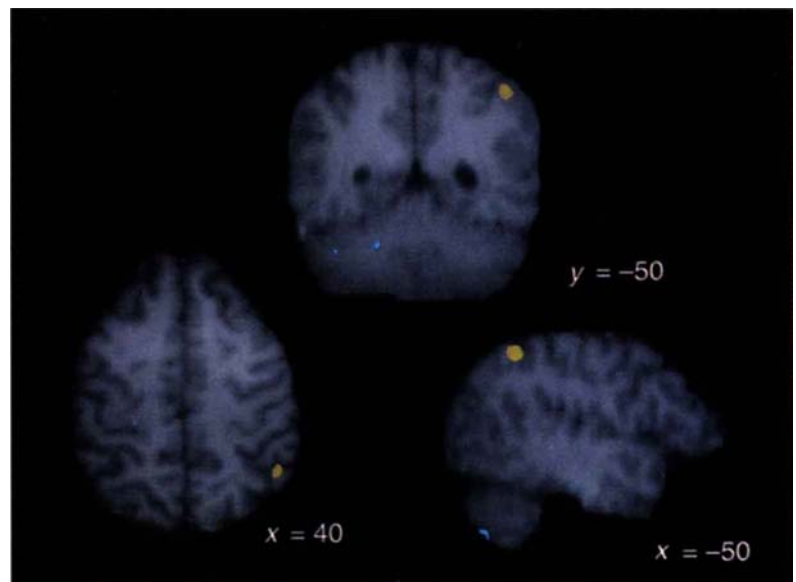


FIG. 1 Time course of prism adaptation as measured by after-effects both in visually targeted reaching and in the perception of hand position. Six subjects were tested separately on an adaptation task where after-effects were measured after 5, 10, 20 and 50 cumulative exposure trials with a rightward displacing prism (group means $\pm$ s.e.m.). Reaching after-effects (open circles) were measured as subjects reached to the remembered location of a visual target with prisms removed. We also measured proprioceptive after-effects (filled circles) as subjects attempted to reach reach with eyes closed to a point in front of the nose. Both types of after-effects (measured as differences between pre- and post-exposure reaching end points) were maximized after 5–10 exposure trials and did not increase further during longer exposures. Units on the ordinate represent screen coordinates along the horizontal axis (the axis of prismatic displacement).

FIG. 2 Coronal (top), horizontal (left), and sagittal (right) views centred on the posterior parietal activation site at Talairach coordinates ($x = -50$, $y = -50$, $z = 40$) for prism adaptation minus error correction difference image, displayed on a single subject's magnetic resonance image (MRI) at a threshold of $P \leq 0.0005$ (raw t value ≥ 4.23 , uncorrected for multiple comparisons). The region of activation within area PEG comprised 38 contiguous pixels that exceeded this threshold (corrected $P = 0.015$). The Talairach plane corresponding to each view is indicated beneath the difference image. Here and in Fig. 3, yellow indicates relative increase in rCBF and blue indicates relative decrease.



ment. Difference images that reflected the net effects of the adaptation process showed selective activation of posterior parietal cortex contralateral to the reaching limb.

The process of adapting to prisms could be mediated by changes at any of several stages in the transformations leading from sensory input to motor output, including alterations in visual perception and the learning of new motor responses^{4,5}. Several lines of evidence indicate that adaptation might be accompanied by alterations in the sense of limb position (or of limb movement), and that this proprioceptive (or kinaesthetic) change may account for much of the alteration in reaching behaviour after exposure to prisms^{6,7}. When prism-adapted subjects are asked to reach directly ahead while their eyes are closed, they show systematic shifts in this proprioceptive measure in the same direction and of similar magnitude to the adaptive shifts observed with visually targeted reaches. In addition, the time courses of these two measures of adaptation are comparable, with subjects displaying maximal post-adaptive after-effects after 5–10 exposure trials (Fig. 1). The after-effects observed in both targeted reaching and the perception of hand position are largely confined to the limb that was active during exposure to prisms^{8,9}, suggesting that the adaptive process does not rely on a global shift in either visual or proprioceptive reference frames. Instead, these observations indicate that targeted reaching depends on accurate correspondence between visual and proprioceptive representations of hand position, and that adaptation to prisms may involve a selective shift in the proprioceptive representation of hand position that brings it back into alignment with the prism-distorted visual input.

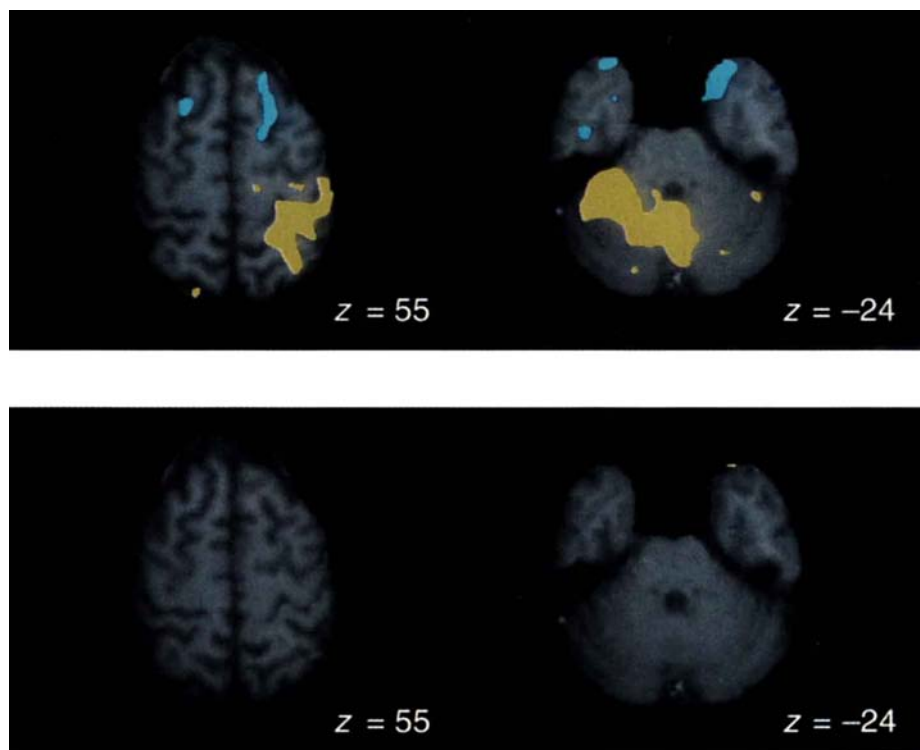
In an effort to localize the cortical regions involved in prism adaptation, we used $H_2^{15}O$ PET to map normalized, relative rCBF in seven right-handed male subjects during three behavioural tasks. In a prism-adaptation task, subjects made visually guided reaches to targets that were viewed through laterally displacing prisms. Because the subjects adapted rapidly to the displacements (Fig. 1), the orientation of the displacing prisms was repeatedly reversed during the scan to maintain the subject in a state of ongoing adaptation. Because the adaptation task also included visual, motor, attentional and error-corrective components, a comparison task was necessary to control for these additional variables. For this purpose, subjects also performed an error-correction task in which the location of the visual target was randomly displaced by 7 cm (either to the left or right) while the subject was in mid-reach, thereby necessitating trajectory corrections that were comparable to those required by the prism-adaptation task. The visual, motor, attentional and error-corrective features of the two tasks were comparable, but as the direction of target displacement was randomized from trial-to-trial and there were

no perceptual distortions, the error-correction task did not require or permit sensorimotor adaptation. Four scans were made during the error-correction task, followed by four scans using the prism-adaptation task. The tasks were performed in separate blocks and in this order to prevent the potential transfer of adaptation effects from prism-adaptation to error-correction scans. To assess the motor components of both tasks, a visual-inspection task was also included, in which the subject passively viewed the visual targets but made no reaches.

Normalized difference images were made that reflected the net relative changes in rCBF between task conditions. The difference images showing prism-adaptation effects corrected for error effects indicated a highly focal region of posterior parietal cortex, contralateral to the reaching hand, that was strongly activated during sensorimotor adaptation to prisms (Fig. 2). The parietal activation site was located on the lateral bank of the intraparietal sulcus, corresponding to area PEG according to a human cytoarchitectonic study¹⁰. Area PEG in humans is a transition zone between the pyramidal cortex of the superior parietal lobule and the granular cortex of the inferior parietal lobule. The comparable area in monkeys lies within the depths of the intraparietal sulcus^{10,11}, a region that contains both separate and superimposed visual and somatosensory maps¹². The selective, unilateral activation of area PEG evident in these images is most likely to reflect the participation of this region in the active process of prism adaptation *per se*. Here we use the term adaptation to refer to the dynamic reorganization that occurs at the central level, in response to the perceived sensory mismatch. However, it is also possible that the activation site reflects the passive registration of the sensory mismatch induced by the prisms. These were the only lateralized behavioural processes that should have been engaged selectively by the prism-adaptation task.

The lateralized activation of area PEG, in the hemisphere contralateral to the reaching limb, was consistent with the well-established principle that prism adaptation effects are largely confined to the prism-exposed limb and do not generalize or transfer to other extremities. Given the anatomical connectivity and physiological characteristics of this region in the monkey, we suggest that the observed activation in the parietal cortex is related to the coordinative remapping that accompanies a prism-induced mismatch between visual and proprioceptive representations of hand position. Psychophysical studies have shown that such a mismatch leads to an adaptive shift in perceived hand position, and through this proprioceptive shift the mismatch between visual and proprioceptive representations is ultimately resolved^{3,5,6}. The time course of the proprioceptive shift directly mirrors the overall adaptive shift, as shown in Fig. 1, indicating

FIG. 3 Comparisons of sensorimotor cortex (left column) and cerebellum (right column) in prism adaptation minus visual inspection (top row), and prism adaptation minus error correction (bottom row), both displayed on a single subject's MRI at a threshold of $P \leq 0.0005$, as in Fig. 2 (raw t value ≥ 4.23 , uncorrected for multiple comparisons). In addition to the prism-adaptation and error-correction scans, two scans were collected during a visual-inspection condition, which required subjects simply to observe the target presentation while alternating the viewing eye, but to make no reaching movements. The targets were stationary in the visual-inspection condition, as in the prism-adaptation task. As this comparison demonstrates, there were substantial movement-related activations in sensorimotor cortex and cerebellum associated with the prism-adaptation task, but these were successfully controlled for by the error-correction task. Note that the horizontal Talairach plane passing through sensorimotor cortex ($z = 55$) is located substantially above the PEG activation site ($z = 40$) shown in Fig. 2. Areas of relatively increased rCBF are shown in yellow, areas of relatively decreased rCBF in blue.



that the progressive elimination of the sensory mismatch and the process of adaptation are inextricably linked, and may not be experimentally dissociable by behavioural means alone.

Surprisingly, we found no activation of the cerebellum in the prism-adaptation minus error-correction images, in spite of the evidence for the involvement of this structure in sensorimotor adaptation^{13,14}. It is possible that cerebellar participation may be limited to the process of error correction that typically accompanies prism adaptation, a mechanism that could be anatomically as well as functionally distinct from the coordinative remapping between visual and proprioceptive representations.

Other lateralized behavioural processes, specifically motor and somatosensory responses, were designed to be evoked in a comparable way by both the prism-adaptation and error-correction tasks and therefore to cancel out in the difference images. Indeed, although there were robust movement-related activations in sensorimotor cortex and cerebellum in both the prism-adaptation minus visual-inspection images (Fig. 3a) and in the error-correction minus visual-inspection images (not shown), none of these effects were evident in the prism adaptation minus error-correction images (Fig. 3b). Although the error-correction task involved shifting spatial attention as the target location moved, it is doubtful that either visual, oculomotor or attentional processes could have explained the unilateral activation of area PEG that was evident in the prism adaptation minus error correction images. This is because, by design, the visual, oculomotor and attentional demands of these two tasks were bilaterally symmetrical, and because any activation related to attention shifts in the error correction task should have appeared as a relative decrease in rCBF in the prism adaptation minus error correction images. To address the issue of possible temporal effects associated with the blocked design, an omnibus F test across repetitions was done for the four scans of each block to determine whether there was any effect of order. At the PEG activation site, no order effects were found within either the error-correction or the prism-adaptation blocks ($P > 0.05$). This indicates that time-dependent factors, such as skill acquisition (a potential consideration as both tasks involved error correction), would be unlikely to explain the observed differences in rCBF across the two conditions.

Whereas a number of previous PET studies have shown activation of posterior parietal cortex in conjunction with various behavioural processes, including visually guided reaching^{15,16}, movement selection^{17,18}, implicit motor imagery^{19,20} and visuospatial attention^{21–23}, none of the parietal areas activated in those studies corresponded precisely to the PEG activation site observed with our prism-adaptation task. We compared the Talairach coordinates²⁴ reported in each of those previous studies with the coordinates of our PEG activation site: in each case, the respective centroids differed by at least 10 mm along one or more of the three Talairach axes. In general, the parietal sites activated in previous studies were located either rostral or medial to area PEG.

Patients with lateralized lesions of area PEG show impairments of hand–eye coordination (termed optic ataxia) when they attempt to make visually guided reaches with the opposite hand²⁵. The reaching inaccuracies exhibited by such patients are even more pronounced when they attempt to reach to remembered targets, without the benefit of visual guidance, even though they may not show impairments with conventional passive tests of proprioception. Taken together with our present results, these observations suggest that area PEG may be important not only in the coordinative remapping that accompanies prism adaptation, but also in the presumably continuous adaptive process that is necessary for maintaining an accurate correspondence between visual and proprioceptive representations of hand position, as required for precise hand–eye coordination. □

Methods

Prism-adaptation task. Subjects wore goggles that positioned a 30 diopter displacing prism over each eye, one oriented to displace the visual field 17° to the left and the other 17° to the right. Subjects were required to reach with the right index finger to one of seven horizontally aligned visual targets that were presented on a touch-screen monitor in a balanced but unpredictable sequence. They were instructed to reach to the targets as they appeared on the screen. Because of the prismatic displacements, reaches included trajectory corrections that began in mid-reach, after the reaching hand entered the subject's field of view. Between each reach, the hand was returned to a resting position on the chest until the next target appeared. To maintain the subjects in a semi-constant state of ongoing adaptation, they alternated the eye through which they viewed

the target every four reaches, effectively switching the direction of the prismatic displacement and preventing complete adaptation to either orientation.

Error-correction task. Goggles contained clear lenses rather than prisms and subjects alternated the viewing eye as they had in the adaptation condition. As a subject began his reach, however, the target location suddenly shifted by 7 cm either to the right or to the left of the initial position, thereby requiring the subject to make a trajectory correction during the reach. The magnitude of the target shift was roughly the same as the visual displacement produced by the prisms, and resulted in reaching movements with the same curved hand paths that were observed in the adaptation task. Using an analysis of variance, which effectively subtracted the error-correction from the prism-adaptation images, activations relating to limb movements, to visual and sensory conditions, and to error-correction processes were effectively removed, leaving only those activations that were related to the adaptive process itself.

Positron-emission tomography. Scans were done with a Siemens ECAT EXACT scanner with septa retracted to allow three-dimensional imaging. A series of 90-s scans were made from each subject, using bolus intravenous injections of H_2^{15}O (25 mCi) that were delivered before the start of each scan. Performance of the designated task began at the same time as the tracer injection, and the scanning period began 10 s after injection. The subject completed 21 or 22 reaches during each 90-s scanning period. High-resolution magnetic resonance images of each subject's brain were obtained in separate sessions (2.5 mm contiguous T1-weighted transaxial spin-echo images (256 × 256 matrix size, time of repetition (TR) = 650 ms, time of echo (TE) = 20 ms) performed on a Philips NT scanner) for individual coregistration with the PET images^{26,27}. PET images were reconstructed with a ramp filter and then smoothed in three dimensions using a Hanning filter with a cut-off frequency of 1 cycle per cm. The final isotropic resolution of the filtered images was 11.2 mm full width at half maximum height. To permit intersubject comparisons, each subject's coregistered PET images were then mapped into the standard Talairach coordinate space. Separate three-way ANOVAs were done (with factors being task, subject and replication), and then contrast analyses were used to produce *t*-statistic images based on comparisons across pairs of conditions (for example, prism adaptation minus error correction)^{28,29}. To minimize type I errors, the search region was restricted to functionally plausible structures, including contralateral sensorimotor, premotor and posterior parietal cortex, and ipsilateral cerebellum. Areas of focal activation that were detected on the *t*-thresholded images were subjected to an analysis of cluster-based significance to calculate corrected *P* values for each activation site³⁰. This analysis was sensitive to the extent as well as the intensity of activation, and corrected for the fact that multiple *t* statistics were generated within the area of search.

Received 7 June; accepted 22 August 1996.

1. Kalaska, J. F. & Crammond, D. J. *Science* **255**, 1517–1523 (1992).
2. Georgopoulos, A. P. *Annu. Rev. Neurosci.* **14**, 361–377 (1991).
3. Redding, G. M. & Wallace, B. *Percept. Psychophys.* **44**, 59–68 (1988).
4. Welch, R. B., Choe, C. S. & Heinrich, D. R. *J. Exp. Psychol.* **103**, 700–705 (1974).
5. Redding, G. M., Clark, S. E. & Wallace, B. *Cogn. Psychol.* **17**, 1–25 (1985).
6. Wilkinson, D. A. *J. Exp. Psychol.* **89**, 250–257 (1971).
7. Baily, J. S. Q. *J. Exp. Psychol.* **24**, 8–20 (1972).
8. Choe, C. S. & Welch, R. B. *J. Exp. Psychol.* **102**, 1076–1084 (1974).
9. Beckett, P. A. *J. Exp. Psychol.* **6**, 433–444 (1980).
10. Eidelberg, D. & Galaburda, A. M. *Arch. Neurol.* **41**, 843–852 (1984).
11. Seltzer, B. & Pandya, D. N. *Brain Res.* **192**, 339–351 (1980).
12. Colby, C. L., Duhamel, J. R. & Goldberg, M. E. *J. Neurophysiol.* **69**, 902–914 (1993).
13. Thach, W. T., Goodkin, H. P. & Keating, J. G. *Annu. Rev. Neurosci.* **15**, 403–442 (1992).
14. Zeffiro, T. *Hum. Brain Mapp. suppl.* **1**, 333 (1995).
15. Grafton, S. T., Mazziotta, J. C., Woods, R. P. & Phelps, M. E. *Brain* **115**, 565–587 (1992).
16. Kawashima, R., Roland, P. E. & O'Sullivan, B. T. *J. Neurosci.* **14**, 3462–3474 (1994).
17. Deiber, M. P. et al. *Exp. Brain Res.* **84**, 393–402 (1991).
18. Stephan, K. M. et al. *J. Neurophysiol.* **73**, 373–386 (1995).
19. Bonda, E., Petrides, M., Frey, S. & Evans, A. *Proc. Natl Acad. Sci. USA* **92**, 11180–11184 (1995).
20. Parsons, L. M. et al. *Nature* **375**, 54–58 (1995).
21. Corbetta, M., Miezin, F. M., Shulman, G. L. & Petersen, S. E. *J. Neurosci.* **13**, 1202–1226 (1993).
22. Corbetta, M., Shulman, G. L., Miezin, F. M. & Petersen, S. E. *Science* **270**, 802–805 (1995).
23. Pardo, J. V., Fox, P. T. & Raichle, M. E. *Nature* **349**, 61–64 (1991).
24. Talairach, J. & Tournoux, P. *Co-Planar Stereotaxic Atlas of the Human Brain* (Thieme, New York, 1988).
25. Perenin, M. & Vighetto, A. *Brain* **111**, 643–674 (1988).
26. Woods, R. P., Cherry, S. R. & Mazziotta, J. C. *J. Comput. Assist. Tomogr.* **16**, 620–633 (1992).
27. Woods, R. P., Mazziotta, J. C. & Cherry, S. R. *J. Comput. Assist. Tomogr.* **17**, 536–546 (1993).
28. Neter, J., Wasserman, W. & Kutner, M. H. *Applied Linear Statistical Models* (Irwin, Burr Ridge, Illinois, 1990).
29. Woods, R. P., Iacoboni, M., Grafton, S. T. & Mazziotta, J. C. in *Quantification of Brain Function Using PET* (eds Meyers, R., Cunningham, V. & Bailey, D.) 353–358 (Academic, New York, 1996).
30. Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C. & Evans, A. C. *Hum. Brain Mapp.* **1**, 210–220 (1994).

ACKNOWLEDGEMENTS. This work was supported in part by the PET Imaging Center, Emory University School of Medicine, Georgia and by grants from the National Institutes of Health. We thank S. T. Grafton for helpful discussions and for his critical reading of the manuscript.

CORRESPONDENCE and requests for materials should be addressed to G.E.A. (e-mail: gea@neuro.emory.edu).

A mechanism for generation of long-range synchronous fast oscillations in the cortex

Roger D. Traub*†, Miles A. Whittington‡, Ian M. Stanford§ & John G. R. Jefferys§

* IBM Research Division, T. J. Watson Research Center, Yorktown Heights, New York 10598, USA

† Department of Neurology, Columbia University, New York, New York 10032, USA

‡ Department of Physiology and Biophysics, Imperial College School of Medicine at St Mary's, London W2 1PG, UK

§ Department of Physiology, The Medical School, University of Birmingham, Birmingham B15 2TT, UK

SYNCHRONOUS neuronal oscillations in the 30–70 Hz range, known as gamma oscillations, occur in the cortex of many species^{1–6}. This synchronization can occur over large distances, and in some cases over multiple cortical areas^{7,8} and in both hemispheres²; it has been proposed to underlie the binding of several features into a single perceptual entity⁴. The mechanism by which coherent oscillations are generated remains unclear, because they often show zero or near-zero phase lags over long distances, whereas much greater phase lags would be expected from the slow speed of axonal conduction. We have previously shown that interneuron networks alone can generate gamma oscillations^{9,10}; here we propose a simple model to explain how an interconnected chain of such networks can generate coherent oscillations. The model incorporates known properties of excitatory pyramidal cells and inhibitory interneurons; it predicts that when excitation of interneurons reaches a level sufficient to induce pairs of spikes in rapid succession (spike doublets), the network will generate gamma oscillations that are synchronized on a millisecond time-scale from one end of the chain to the other. We show that in rat hippocampal slices interneurons do indeed fire spike doublets under conditions in which gamma oscillations are synchronized over several millimetres, whereas they fire single spikes under other conditions. Thus, known properties of neurons and local synaptic circuits can account for tightly synchronized oscillations in large neuronal ensembles.

In visual cortex, gamma oscillations can be evoked by visual stimulation. When the stimulus is topologically connected (a long bar, for example), the oscillations can be coherent over distances of up to 7 mm, with zero or near-zero (<3 ms) average phase lag¹. Coherence of in-phase gamma oscillations has also been observed between primary and associational visual cortices^{7,8}, and across the corpus callosum², where antidromic axonal conduction delays are estimated to be 2.73 ± 2.38 ms (ref. 11). In humans, gamma oscillations are observed in magnetic signals emanating from large regions of cerebral cortex and subcortical structures¹². The mechanism of coherence is not obvious, as axonal conduction in local circuit axonal processes is slow (~ 0.5 m s⁻¹, or 1 cm in 20 ms, for hippocampal Schaffer collaterals¹³; 0.15 – 0.55 m s⁻¹ for neocortical pyramidal cell collaterals¹⁴; and 0.06 – 0.2 m s⁻¹ for neocortical inhibitory collaterals¹⁵).

The coherent oscillation problem can be broken down into two questions: how do local circuits, or even individual cells¹⁶, generate gamma-frequency oscillations, and what happens when oscillating local circuits are synaptically interconnected? We begin with a known local-circuit interneuron network gamma oscillator^{9,10}. We then use simulations to review how this local circuit behaves when both pyramidal cells and interneurons are present, and go on to demonstrate how synaptically interconnected local circuits can produce in-phase oscillations over significant distances. In our model, synchronized oscillations are apparent in single trials (unlike the model of ref. 17), even when conduction delays total 20 ms from one end of a chain of oscilla-

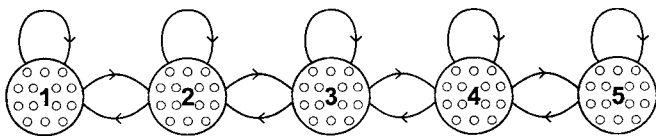


FIG. 1 Structure of the model. There are 5 cell groups, each with 8 pyramidal cells (e-cells, simulated as in ref. 26) and 8 basket cells (i-cells, simulated as in ref. 18). There are 4 types of synaptic connections: $e \rightarrow e$ (in some simulations), $e \rightarrow i$, $i \rightarrow e$, $i \rightarrow i$. Connections occur between cells in each given group and between neighbouring groups, so that each cell has 16 excitatory and 16 inhibitory inputs. Inhibitory connections were GABA_A, with IPSCs of abrupt onset and exponential decay ($\tau = 10$ ms). Excitatory connections were AMPA-receptor-mediated (conductance proportional to a 2-ms α -function between pyramidal cells, a 1-ms α function on i-cells) and NMDA-receptor-mediated ($\tau = 150$ ms between pyramidal cells, 60 ms on i-cells). Peak unitary inhibitory conductances were 4 nS on e-cells, 1.5 nS on i-cells. Axon conduction delays within a group are considered negligible, whereas between groups they could range from 0 to 10 ms; in most simulations, intergroup conduction delays were 4 or 5 ms. Interneurons were biased with a small hyperpolarizing current to suppress spontaneous firing. Pyramidal cells received a somatic driving current of $1.6 + x$ nA, where x was varied over a 0.2 nA range across each group; there was an additional 0.01 nA spread across the cells, so that no 2 cells received identical currents. With such a current, isolated pyramidal cells would fire a single burst and then settle into a repetitive firing mode^{26,27}. Data were collected for the last 750 ms of 1,500-ms runs. Noise was included in the form of ectopic action potentials in e-cell and i-cell axons. Simulations were run on IBM SP1 and SP2 parallel computers; a 1,500-ms simulation took 1.8 h on the SP2.

tors to the other. Finally, we provide experimental evidence for two central predictions of our model.

Local circuits of inhibitory interneurons—in hippocampus and neocortical layers 2/3 of the rat—can generate gamma-frequency oscillations during pharmacological blockade of ionotropic glutamate receptors, when the cells are excited by metabotropic glutamate receptors and inhibit each other by GABA_A receptors⁹. A circuit model, using a 51-compartment axon/soma/dendritic model for each interneuron¹⁸, accounts for the observed properties of the network oscillation, including the dependence of network frequency on unitary inhibitory postsynaptic conductance and time course. The model works when a sufficiently high synaptic connectivity exists^{10,19}. This model also correctly predicts a break-up of the network oscillation when a sufficiently low frequency is attained¹⁰. In simulations, eight interneurons are sufficient to generate synchronized gamma oscillations, when all-to-all connectivity exists. In such small local circuits, axon conduction delays are probably negligible.

The behaviour of a local circuit generator of gamma oscillations was examined when pyramidal cells were included, with interneurons, using the following synaptic interactions: in the excitatory connection from pyramidal cells to interneurons, both slow (for example, NMDA-receptor-mediated) and fast (for example, AMPA-receptor-mediated) synaptic actions; and in the inhibitory connection from interneurons to other interneurons and to pyramidal cells, GABA_A-receptor-mediated synaptic action. Axon conduction delays were <0.5 ms for i-cells (inhibitory cells) and e-cells (excitatory or pyramidal cells) (R.D.T., J.G.R.J. and M.A.W., unpublished results). In this model, pyramidal cells and interneurons fired with near-zero phase lag on average, and on occasion the fast excitatory postsynaptic potentials (EPSPs) to interneurons induced doublet firing, at about 5-ms intervals, with the second interneuron spike usually following the pyramidal cell spike (see also Fig. 2b). (In the model, doublet firing did not occur as an intrinsic property of the interneurons.) This property of doublet firing suggested a way that oscillators might be connected together with conduction delays and still oscillate coherently with near-zero phase lag.

Figure 1 illustrates the structure of the model: a chain of five cell groups, each group containing eight pyramidal cells and eight

interneurons, with all-to-all intragroup connectivity, and negligible intragroup axon-conduction delays. Cells in each group also receive excitatory and inhibitory input from one-half or all of the excitatory and inhibitory neurons in each neighbouring group, so that all cells have the same number of inputs. Conduction times for the intergroup excitatory and inhibitory axons were treated as free parameters and were manipulated independently (from 0 to 10 ms).

Figure 2a illustrates the behaviour of the network model when intergroup synaptic connections are cut. Individual cell groups oscillate at about 40 Hz, as seen in the autocorrelation, performed on the average of four pyramidal cell voltages from group 1, over a period of 750 ms. The flat intergroup cross-correlation shows that oscillations in the different ends of the chain are, on average, independent and uncorrelated; thus, there is no artefactual synchrony, as might exist in a noiseless system with identical stimuli to each group. Interneurons usually fire single action potentials, but spike doublets, synaptically induced, occur on occasion. Note that axo-axonic interneurons readily fire spike doublets during intracellular current injection²⁰, and it is not unusual for other interneurons, when sufficiently excited, to fire action potentials at short intervals¹⁹.

In contrast (Fig. 2b), when intergroup synaptic connections are present, with 4-ms axon conduction delays between neighbouring

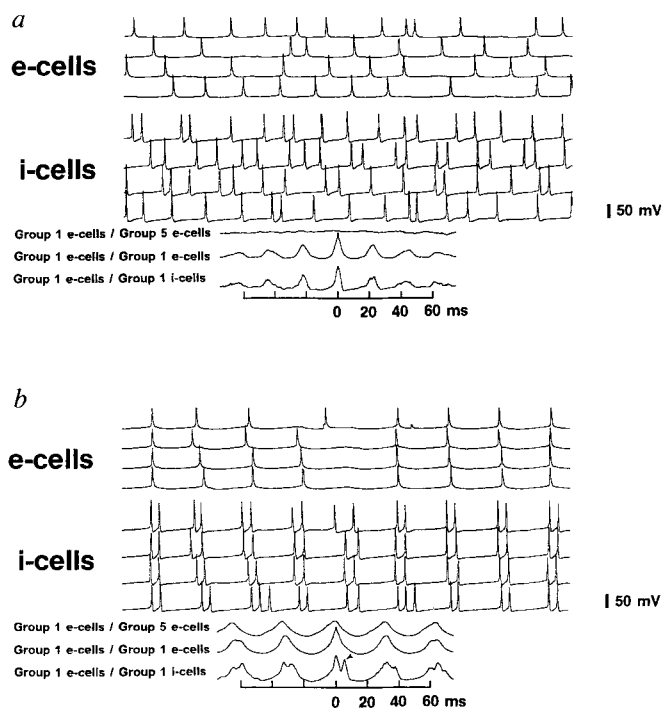
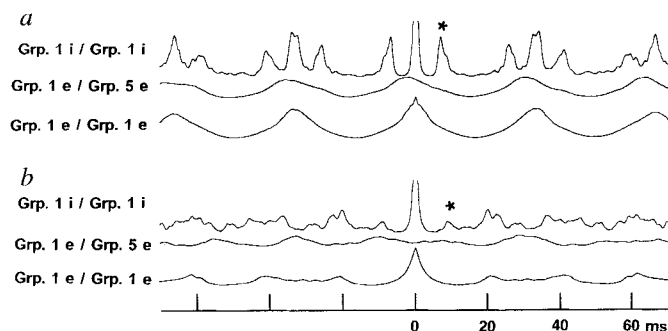


FIG. 2 Tight long-range synchronization exists in the model even with intergroup axon conduction delays. *a*, The 5-cell groups are uncoupled from each other. Potentials of 4 e-cells (top) and 4 i-cells (bottom) are shown; cells respectively from groups 1, 2, 4, 5. The autocorrelation and cross-correlation were obtained by averaging 4 e-cell potentials from groups 1 and 5. The group 1 autocorrelation reveals a peak at 25 ms (40 Hz), but the group 1/group 5 cross-correlation is flat. The group 1 e-cell/i-cell cross-correlation indicates that, on average, nearby e-cells and i-cells fire in phase. *b*, Same parameters as in *a*, but intergroup connections exist with axon conduction delays of 4 ms ($e \rightarrow i$, $i \rightarrow e$, $i \rightarrow i$). The network frequency slows to 32 Hz, and tight correlations exist between groups 1 and 5 (the cross-correlation peak is at -0.9 ms). Note that now the interneurons fire mostly doublets at average 5.25-ms intervals. The group 1 e-cell/i-cell cross-correlation indicates that the first interneuron spike occurs with the same phase, on average, as the e-cell spike; doublet firing appears as the second peak, lagging the first peak by 5.25 ms; $e \rightarrow e$ connections were blocked in both simulations.



groups (hence 16 ms total delay between ends of the chain), coherent cellular oscillations occur throughout the system. The mean phase lag between opposite ends of the model chain is only 0.9 ms. The oscillation frequency also slows from 40 to 32 Hz. Note that in the simulation shown in Fig. 2b, the interneurons almost always fire in doublets or occasionally triplets.

To show that interneuron doublet firing provides the 'glue' that allows long-range synchronization to occur, we performed a manipulation to suppress doublet firing. In Fig. 3, the simulation of Fig. 2 was repeated twice, with progressively smaller values of the unitary $e \rightarrow i$ AMPA receptor conductance. When this conductance is still large enough for the second interneuron spike to be recruited, so for interneuron doublet firing to occur (Fig. 3a), strong correlations in cellular oscillations still occur from one end of the array to the other. As doublet firing is reduced sufficiently, long-range correlations in the oscillation are attenuated.

We next sought experimental evidence for interneuron doublets. Tetanic stimulation of the CA1 region of rat hippocampal slices is known to induce gamma-frequency synaptic potentials in pyramidal cells that are blocked by bicuculline and hence are probably generated by networks of inhibitory cells⁹. We found (Fig. 4) that simultaneous tetanic stimulation of two CA1 sites, at distances up to 4 mm apart, induces gamma-frequency field potential

FIG. 3 Interneuron doublet firing is required for long-range synchronization in the model. *a*, When the $e \rightarrow i$ AMPA conductance is reduced to 50% of the value in Fig. 2, synchronized oscillations still occur (29 Hz), although now with a 2.5-ms lag in the e-cell cross-correlation (group 1/group 5). Interneuron doublet firing is indicated by the peaks (asterisk) around the (truncated) central peak in the interneuron autocorrelation (based on the average of 4-i-cell potentials from group 1). *b*, When the $e \rightarrow i$ AMPA conductance is further reduced to 24% of the value in Fig. 2, synchronized oscillations continue within individual cell groups (note the small but distinct 20 ms/50 Hz peak in the e-cell autocorrelation), but between-group correlations are lost. At the same time, doublet firing in interneurons is suppressed (asterisks in interneuron autocorrelation).

oscillations that are coherent (<1 ms phase lag in the cross-correlation). Physiologically identified stratum pyramidale interneurons fire rhythmic action potential singlets after local tetanic stimulation, whereas they usually fire doublets following double simultaneous tetanic stimulation (Fig. 4). Doublet firing was observed in 4 of 6 interneurons under such conditions. The oscillation frequency is reduced after the dual tetanic stimulation, as predicted by the model (Fig. 2). A similar reduction in gamma

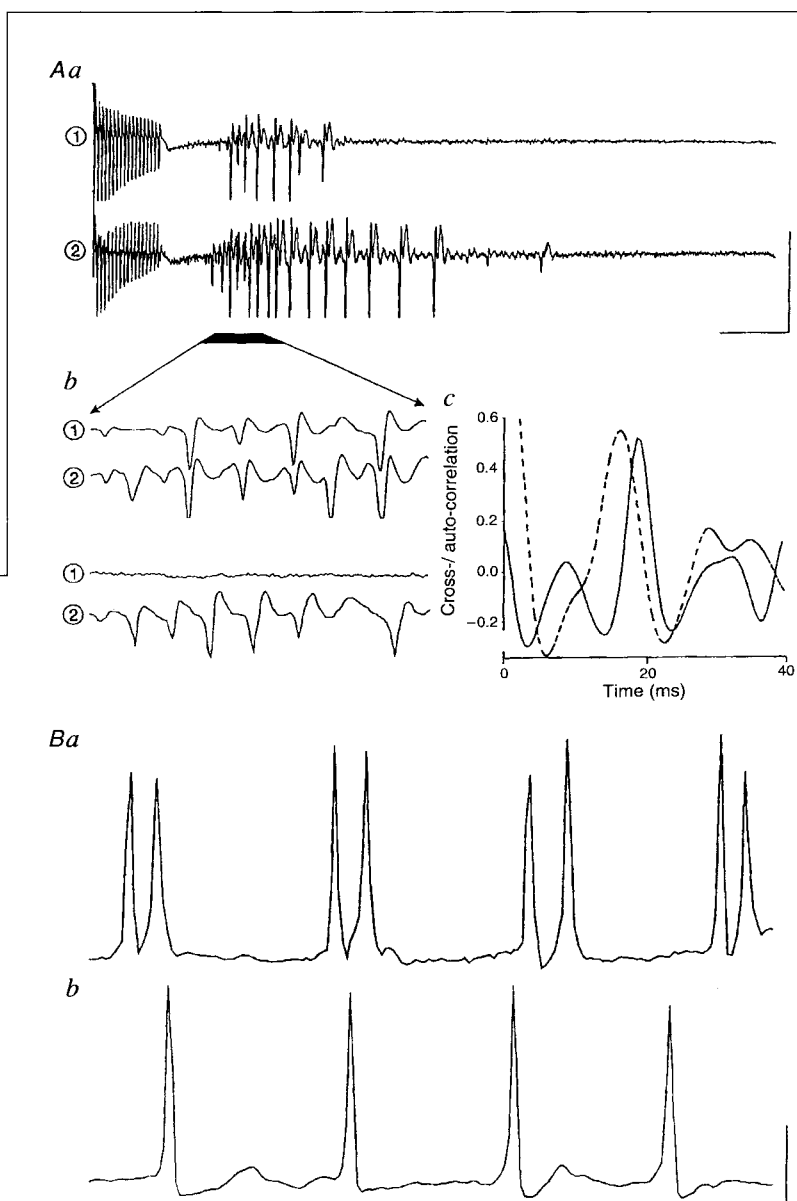


FIG. 4 Interneuron spike doublets occur under conditions associated with long-range synchrony in the hippocampal slice. *A*, Recordings of extracellular population potentials for two sites (1 and 2) in the hippocampal CA1 subfield separated by 4 mm. *a*, Trains of 20 stimuli sufficient to generate a half-maximal population spike were delivered at 100 Hz by bipolar electrodes placed at the border of stratum oriens/alveus at the CA2 (site 1) and subicular (site 2) ends of CA1. 50 to 150 ms following tetanus, rhythmic oscillating field potentials were seen in stratum pyramidale at both sites. Scale bar: 200 ms, 10 mV. *b*, Analysis of trains of oscillations at 2 sites revealed a highly synchronous organization of population spikes, with cross-correlation analysis showing a frequency of 52 Hz (cross-correlation, solid line). *c*, Single tetanic stimulation at site 2 also generated an oscillatory field potential locally, with a faster frequency of population spikes, 62 Hz (autocorrelation, broken line). Timescale was expended 7-fold for example traces. *B*, *a*, Recordings from electrophysiologically identified stratum pyramidale interneurons after paired tetanic stimulation revealed rhythmic doublet firing. *b*, Single-site tetanus generated interneuron oscillations locally with no doublet firing. Scale bar: 20 ms, 40 mV.

frequency oscillation with increasing size of the participating neuronal ensemble has been observed *in vivo*²¹.

How might interneuron doublets permit long-range synchrony? König and Schillen²² showed that global coherent oscillations can occur when a local circuit time constant matches the delay time between neighbouring local circuits. They unrealistically assumed, however, that the delay between pyramidal cell firing and excitation of nearby interneurons should match the conduction delay between separated cell groups. Doublet firing of the interneurons circumvents this need for slow conduction in the local circuit. Instead, the doublet interval of 4–5 ms acts as the local circuit time constant, which can then match approximately the conduction delay between spatially separated circuits. The second interneuron spike in each doublet is stimulated in part by nearby pyramidal cells and in part by the delayed excitation from spatially separated pyramidal cells. Synchronization between two model columns with conduction delays of up to 5 ms has been shown without interneuron doublets, but with interneuron bursts perhaps playing a similar role²³.

Our model predicts (unlike the models of refs 22 and 24) that pyramidal cells and at least some interneuron spikes should occur with near-zero phase lag⁶, and that progressive attenuation of pyramidal cell to interneuron excitation should lead to loss of long-range synchrony before loss of local synchrony (Fig. 3). In the neocortex, large basket cells in layer III have relatively long axons contacting pyramidal cells as well as other large basket cells²⁵. Such cells might allow our proposed mechanism to underlie long-range synchrony of '40 Hz' oscillations *in vivo*. □

Methods

Transverse 450- μ m-thick hippocampal slices^{9,10} were prepared from adult male Sprague-Dawley rats and maintained at 35 °C at the interface between warm, wetted O₂/CO₂ and artificial cerebrospinal fluid containing (in mM): NaCl 135, KCl 4, NaH₂PO₄ 1.25, CaCl₂ 1.5, MgCl₂ 0.8, glucose 10. Stimuli: 50 μ s, 6–14 V, were delivered via bipolar silver wire electrodes. Interneurons were impaled at the level of stratum pyramidale with glass microelectrodes containing 2 M potassium methylsulphate, resistance 50–70 M Ω . Neurons had resting membrane potentials between –59 and –68 mV and input resistances of 35–45 M Ω . Presumed interneurons fired narrow (base width 2 ms) action potentials with a rapid, monophasic after hyperpolarization, at high frequencies (>400 Hz), with no accommodation on depolarization.

Received 20 June; accepted 26 August 1996.

- Gray, C. M., König, P., Engel, A. K. & Singer, W. *Nature* **338**, 334–337 (1989).
- Engel, A. K., König, P., Kreiter, A. K. & Singer, W. *Science* **252**, 1177–1179 (1991).
- Gray, C. M. *J. Comput. Neurosci.* **1**, 11–38 (1994).
- Singer, W. & Gray, C. M. *Annu. Rev. Neurosci.* **18**, 555–586 (1995).
- Soltesz, I. & Deschênes, M. *J. Neurophysiol.* **70**, 97–116 (1993).
- Bragin, A., Jandó, G., Nádasdy, Z., Hetke, J., Wise, K. & Buzsáki, G. *J. Neurosci.* **15**, 47–60 (1995).
- Engel, A. K., Kreiter, A. K., König, P. & Singer, W. *Proc. Natl Acad. Sci. USA* **88**, 6048–6052 (1991).
- Frien, A., Eckhorn, R., Bauer, R., Woelbern, T. & Kehr, H. *Neuroreport* **5**, 2273–2277 (1994).
- Whittington, M. A., Traub, R. D. & Jefferys, J. G. R. *Nature* **373**, 612–615 (1995).
- Traub, R. D., Whittington, M. A., Colling, S. B., Buzsáki, G. & Jefferys, J. G. R. *J. Physiol.* **493**, 471–484 (1996).
- Innocenti, G. M. *Arch. Ital. Biol.* **118**, 124–188 (1980).
- Ribary, U. et al. *Proc. Natl Acad. Sci. USA* **88**, 11037–11041 (1991).
- Andersen, P., Silfverius, H., Sundberg, S., Sveen, O. & Wigström, H. *Brain Res.* **144**, 11–18 (1978).
- Murakoshi, T., Guo, J.-Z. & Ichinose, T. *Neurosci. Lett.* **163**, 211–214 (1993).
- Salin, P. A. & Prince, D. A. *J. Neurophysiol.* **75**, 1589–1600 (1996).
- Linäs, R. R., Grace, A. A. & Yarom, Y. *Proc. Natl Acad. Sci. USA* **88**, 897–901 (1991).
- Wilson, M. A. & Bower, J. M. *Neural Comput.* **3**, 498–509 (1991).
- Traub, R. D. & Miles, R. J. *Comput. Neurosci.* **2**, 291–298 (1995).
- Sik, A., Penttonen, M., Ylinen, A. & Buzsáki, G. *J. Neurosci.* **15**, 6651–6665 (1995).
- Buhl, E. H. et al. *J. Neurophysiol.* **71**, 1289–1307 (1994).
- Bauer, R., Brosch, M. & Eckhorn, R. *Brain Res.* **669**, 291–297 (1995).
- König, P. & Schillen, T. B. *Neural Comput.* **3**, 155–166 (1991).
- Bush, P. & Sejnowski, T. J. *Comput. Neurosci.* **3**, 91–110 (1996).
- Eckman, F. H. & Freeman, W. J. *Brain Res.* **528**, 238–244 (1990).
- Kisvárdy, Z. F., Beaulieu, C. & Eysel, U. T. *J. Comp. Neurol.* **327**, 398–415 (1993).
- Traub, R. D., Jefferys, J. G. R., Miles, R., Whittington, M. A. & Tóth, K. J. *Physiol.* **481**, 79–95 (1994).
- Wong, R. K. S. & Prince, D. A. *J. Neurophysiol.* **45**, 86–97 (1981).

ACKNOWLEDGEMENTS. Supported by IBM, the Wellcome Trust, and the Human Frontier Science Program. We thank H. Monyer, R. K. S. Wong, R. Miles, G. Buzsáki, N. Kopell and R. Linsker for helpful discussion, and R. Walkup and J. Jann for advice with parallel computers.

CORRESPONDENCE and requests for materials should be addressed to R.D.T. (e-mail: traub@watson.ibm.com).

Differentiation of adult hippocampus-derived progenitors into olfactory neurons *in vivo*

Jaana O. Suhonen*, Daniel A. Peterson, Jasodhara Ray & Fred H. Gage

Laboratory of Genetics, The Salk Institute, 10010 North Torrey Pines Rd, La Jolla, California 92037-1099, USA

NEUROGENESIS continues throughout adulthood in discrete regions. Proliferative zones include the subependymal zone^{1–4}, from where progenitors migrate along the rostral migratory pathway to differentiate into neurons in the olfactory bulb⁴, and the hippocampal subgranular zone, where they migrate and differentiate into granule neurons^{5–7}. Progenitors isolated from adult subependymal zone exhibit *in vitro* neurogenesis when stimulated with epidermal^{8,9} or fibroblast growth factor¹⁰. Cultured adult rat hippocampal progenitors (AHPs) grafted to adult rat hippocampus show site-specific neuronal differentiation¹¹. Here we investigate determinants of multipotentiality in the adult central nervous system, by grafting AHPs into homotypic (hippocampus) or heterotypic (the rostral migratory pathway) neurogenic sites or a heterotypic, non-neurogenic site (the cerebellum). We found that grafts into neurogenic, but not non-neurogenic sites, showed neuronal differentiation. Furthermore, AHPs grafted in the rostral migratory pathway migrated into the olfactory bulb, differentiating into tyrosine-hydroxylase-positive neurons, a non-hippocampus phenotype. These results reveal that AHP populations can respond to persistent neuronal differentiation cues in the adult central nervous system.

AHP cells in culture remain uncommitted, with only a small fraction differentiating into cells with mature neuronal or glial phenotypes^{11,12}. To test whether AHPs cultured for over two years differentiate *in vivo* in response to endogenous local cues, cells labelled with 5-bromo-2'-deoxyuridine (BrdU), or an adenoviral vector carrying the *lacZ* gene, were grafted to adult rat hippocampus, cerebellum and rostral migratory pathway (RMP) (Fig. 1a). Control experiments, in which BrdU- and adenoviral-labelled cells were sonicated or freeze-thawed before grafting, showed no transfer of markers to endogenous cells¹¹. The graft position, distribution of AHPs and their fate were examined at one and eight weeks post-grafting using BrdU- and β -galactosidase immunostaining. Immunostaining revealed that AHPs had survived at all time points in each target zone with no tumour formation (Fig. 1b–g). Of 75,000 AHPs grafted per site, 27,136 $\pm$ 5,924 (mean $\pm$ s.e.m.) were present in the olfactory bulb, 35,863 $\pm$ 1,123 in the hippocampus, and 26,584 $\pm$ 2,323 in the cerebellum 8 weeks after grafting. There was no significant difference (ANOVA, $\alpha = 0.05$) in AHP survival, indicating that graft site does not influence survival.

Examination of cell distribution at one week revealed the dispersion of AHPs in the hippocampus and cerebellum (Fig. 2a). However, AHPs in RMP migrated over 2 mm rostrally along the RMP and subependymal zone (SEZ) (Fig. 2a), where most (73%) cells were observed in SEZ, with some (19%) in the olfactory granule cell layer and few (8%) in the glomerular cell layer (Table 1a). Eight weeks post-grafting, cerebellar AHP distribution did not differ from that at one week (Fig. 2b), suggesting that dispersion rather than specific migration was occurring. In hippocampus, AHPs distributed in the CA1–3 regions and dentate gyrus¹¹. In the olfactory bulb, AHPs migrated

* Present address: Dept of Neurology and Restorative Neurology, Tampere University Hospital, PO Box 2000, FIN-33521 Tampere, Finland.

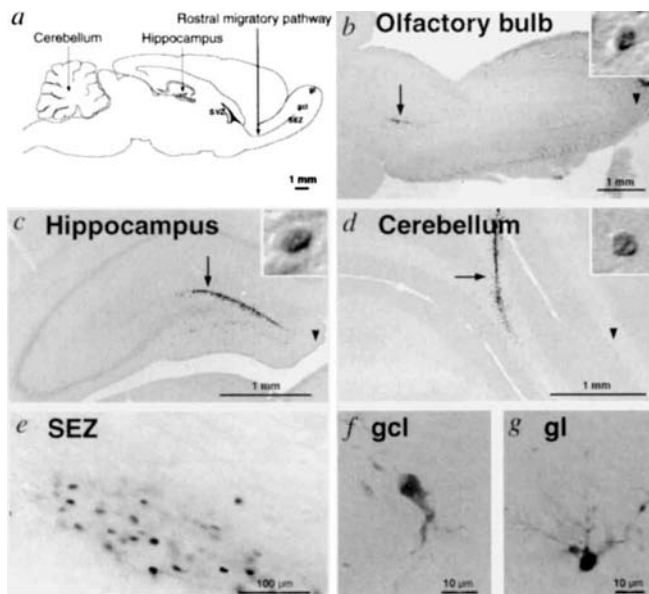


FIG. 1 Migration of AHPs grafted into the adult rat brain. *a*, Parasagittal view of the adult rat brain showing the three graft sites (arrows) with endogenous proliferative regions, the dentate gyrus and SVZ, shaded. BrdU-immunoreactive AHPs at eight weeks post-grafting (arrows) in *b*, the OB/RMP (a heterotypic neurogenic site); *c*, HC (a homotypic neurogenic site); and *d*, CB (a heterotypic non-neurogenic site). Insets in *b*–*d* show the most distant AHPs (arrowheads) from the graft site at higher power. AHPs immunoreactive for β -gal grafted to the RMP distributed along the SEZ (*e*) and in olfactory granule (gcl) (*f*) and glomerular (gl) (*g*) cell layers.

up to 5 mm rostrally (Fig. 2*b*), with only 9% remaining in SEZ, whereas most were found in the olfactory granule (69%) and glomerular (22%) cell layers (Table 1*a*). The extent and pattern of migration were similar to those for olfactory bulb development^{4,13} and in adult neurogenesis^{14–16}, or following homotypic engraftment of freshly dissected cells from adult mice subventricular zone (SVZ; ref. 14). These results indicate that AHPs can respond to guidance cues present in the olfactory system, resulting in long-distance migration and laminar distribution similar to endogenous SVZ-derived progenitors of the host brain.

To investigate the long-distance migration of AHPs and the behaviour of the migrating cells, we examined the expression of polysialic acid neural-cell adhesion molecules (N-CAM; refs 15–18), glial fibrillary acidic protein (GFAP; an astroglia marker), and TuJ1 (a class III β -tubulin), previously used to mark neuroblasts¹³. Although expressed preferentially during embryonic development¹⁹, polysialic acid N-CAM expression continues throughout adulthood in SVZ, olfactory system and hippocampus^{15,16,20,21}. We detected polysialic acid N-CAM in the dentate gyrus, SVZ and olfactory system, but not in the cerebellum. Although AHPs were associated with polysialic acid N-CAM in hippocampus at one week, none was found at eight weeks post-grafting. Many AHPs were associated with polysialic acid N-CAM in chains along the SEZ at one week after grafting, as in endogenous SVZ cells^{15,16} (Fig. 2*c*). By eight weeks, only a few N-CAM-associated AHPs remained within the SEZ, with most located in granule and glomerular cell layers (Fig. 2*d, e*), suggesting radial migration to superficial layers, where neuronal differentiation was observed (Fig. 3*b–d*). In addition to association with polysialic acid N-CAM, support for parallel behaviour of endogenous SVZ progenitors and grafted AHPs comes from the observation that at one week no AHP expressed GFAP in the RMP, although many BrdU-positive cells were TuJ1-positive (Fig. 2*f–h*), indicating that these migratory grafted AHPs in RMP were neuroblasts.

To determine whether AHPs can differentiate into neurons by eight weeks post-grafting, immunofluorescent labelling was per-

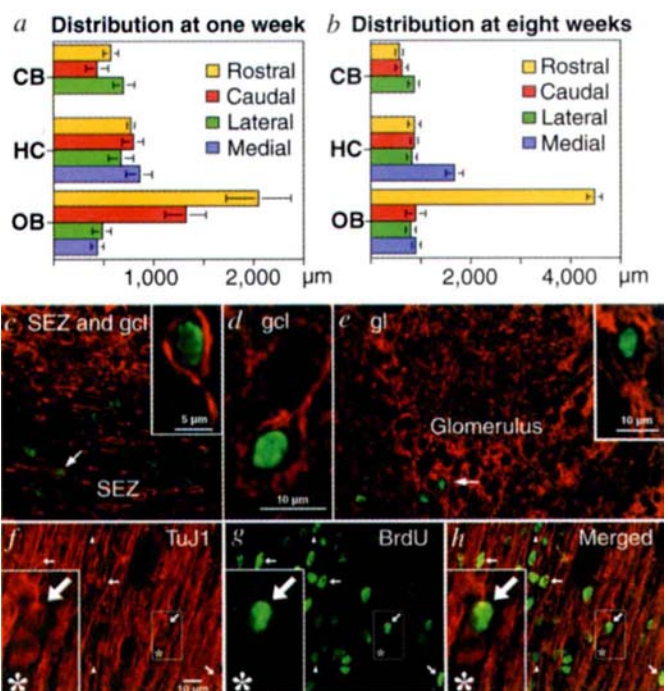


FIG. 2 Maximal distribution of AHPs grafted to CB, HC and OB from the graft centre in four anatomical axes at one (*a*) and eight (*b*) weeks post-grafting ($\pm$ s.e.m.). One week post-grafting, some migrating BrdU-immunoreactive cells expressed association with PSA-N-CAM in the SEZ and gcl (*c*). At eight weeks post-grafting, association of grafted BrdU-immunoreactive cells with PSA-N-CAM was also found in gcl (*d*) and gl (*e*). Cells indicated by the arrows shown at higher power in the insets. Association of TuJ1 with BrdU-immunoreactive cells in RMP (*f–h*). Merged images of cells (*h*) immunostained for TuJ1 (*f*) and BrdU (*g*) showed the cells expressing both markers (arrows), BrdU only (arrowheads) or TuJ1 only (cell above the asterisk). Boxed areas are shown at higher-power magnification as insets at the bottom of *f–h*.

formed with antibodies against tyrosine hydroxylase, calbindin, BrdU, NeuN, and GFAP. In the cerebellum, the heterotypic non-neurogenic site, AHPs did not express neuronal markers (Table 1*b*), suggesting that adult cerebellum may be deficient in neuronal differentiation cues, in contrast to that of the neonate^{22,23}. Poor survival did not preclude differentiation, because AHP survival in cerebellum was equivalent in hippocampus and olfactory bulb. In hippocampus, the homotypic neurogenic site, only AHPs that migrated into neuronal layers of the dentate gyrus expressed the neuronal markers, NeuN (48%) (ref. 24) and calbindin (35%) (ref. 25; Table 1*b*). No tyrosine-hydroxylase-positive cells were detected in the hippocampus (Table 1*b*).

Although olfactory periglomerular neurons are reported to be immunopositive to calbindin and tyrosine hydroxylase^{25,26} and immunonegative to NeuN²⁴, double-labelling revealed that calbindin and tyrosine hydroxylase are expressed in separate, non-overlapping periglomerular subpopulations (Fig. 3*b, c*), whereas olfactory granule cells are NeuN-positive but calbindin-negative. When grafted to RMP, the heterotypic neurogenic site, 10% of AHPs within the glomeruli expressed calbindin (Fig. 3*b* and Table 1*b*), and 16% of AHPs within the granule cell layer were NeuN-positive (Fig. 3*d* and Table 1*b*). Furthermore, 7% of AHPs in the glomeruli expressed tyrosine hydroxylase (Fig. 3*c* and Table 1*b*), a marker for dopaminergic neurons. Tyrosine-hydroxylase-positive neurons are not found in hippocampus or in AHPs in culture, or in AHPs grafted to the hippocampus (data not shown). Expression of tyrosine hydroxylase indicates that the AHP population may differentiate in a multipotential fashion after grafting in response to specific local cues. AHPs immunostaining for β -galactosidase displayed an interneuron-like morphology in granule and glomerular

TABLE 1 Distribution of grafted cells

(a) Regional distribution of AHPs in olfactory bulb								
	SEZ	Granule cell layer	Glomeruli	Total				
1 week	258* (73%)	67 (19%)	29 (8%)	354 (100%)				
8 weeks	83 (9%)	611 (69%)	201 (22%)	895 (100%)				
(b) Phenotypic distribution of AHPs in all grafted areas								
	Olfactory bulb			Hippocampus		Cerebellum		
	Glomeruli	Granule cell layer	SEZ	Granule cell layer	Area CA1	Area CA3	Granule cell layer	Purkinje cell layer
BrdU ⁺	203†	204	205	102	105	101	100	50
TH ⁺	7%‡	0%	0%	0%	0%	0%	0%	0%
Calb ⁺	10%	0%	2%	48%	3%	4%	0%	0%
BrdU ⁺	201†	202	204	100	103	102	102	60
NeuN ⁺	0%‡	16%	0%	35%	0%	0%	0%	0%
GFAP ⁺	21%	25%	29%	4%	34%	31%	28%	17%

a, Grafted cells at one and eight weeks after implantation into the rostral tip of RMP were counted from BrdU-immunoreactive sections in the same olfactory bulb regions used for data collection in b (one section per animal; five animals per time point). b, Eight weeks post-grafting, sections including olfactory bulb, hippocampus and cerebellum were triple-labelled with neuronal markers (tyrosine hydroxylase, TH; calbindin, Calb; and NeuN) and an astrocytic marker (GFAP) and analysed by confocal microscopy. BrdU-immunoreactive cells were counted in regions outside the injection site in each target zone. Because of the high abundance of BrdU-positive cells within neuronal layers, an upper limit of 200 sampled cells was imposed. SEZ, subependymal zone.

* The total number of BrdU⁺ cells counted per area.

† The total number of BrdU⁺ cells counted per area.

‡ Percentage of cells double-labelled for BrdU and the indicated marker.

cell layers (Fig. 1f, g). The observations that at one week grafted AHPs in RMP are GFAP-negative but TuJ1-positive, and associated with polysialic acid N-CAM, and that at eight weeks the grafted AHPs in the olfactory bulb express tyrosine hydroxylase, calbindin and NeuN, indicate that AHPs can behave as endogenous progenitors.

Previous studies have used fetal or neonatal tissue to investigate the regional environmental cues underlying neuronal differentiation^{23,27,28}, but our results indicate that the adult nervous system contains progenitor cells that can be expanded *in vitro* and can respond to persistent cues in adult brain for survival, migration and neuronal differentiation. Furthermore, the existence of AHPs positive for tyrosine hydroxylase in the olfactory bulb suggests that regional cues in the adult central nervous system (CNS) can direct differentiating neurons down specific phenotypic pathways. Our observations also suggest that regional cues within the adult CNS may direct phenotypic differentiation. Identifying the conditions for phenotypic differentiation of adult CNS progenitors may provide new strategies for treatment of neurodegenerative diseases and trauma. □

Methods

Grafting of AHPs in the brain. AHPs (75,000 cells in 1.5 µl) labelled with BrdU and/or infected with an adenovirus vector expressing cytoplasmic *E. coli lacZ* (ref. 11) were stereotactically injected into the rostral tip of RMP (AP +4.8, ML ±2.0, DV -3.0 from dura), cerebellum (CB) (AP -10.4, ML ±2.0, DV -4.0 from dura) and hippocampus (HC) (AP -3.5, ML ±3.0, DV -3.9 from skull, with nose bar at 5 mm up) of adult female Fischer-344 rats. (AP, Anteroposterior axis; ML, mediolateral axis; DV, dorsoventral axis.)

Immunostaining and analysis of brain tissues. At one (short-term; *n* = 10) and eight (long-term; *n* = 10) weeks post-grafting, animals were perfused (4% paraformaldehyde), the brains sectioned (40 µm), and immunoperoxidase-stained for BrdU and β-galactosidase (β-gal)¹¹. For immunofluorescent staining, sections were pretreated for BrdU detection and stained¹¹ with mouse anti-NeuN (1:10; from R. Mullen²⁴), rabbit antibody against calbindin-D (28K) (1:1,000; Swant), rabbit anti-GFAP (1:250; Chemicon), mouse antibody against tyrosine hydroxylase (TH) (1:250; Boehringer Mannheim), mouse anti-MenB (to detect polysialic acid (PSA)-N-CAM residues; 1:1,000 with biotin-streptavidin amplification; from G. Rougon²⁹), mouse anti-TuJ1 (1:1,000; from A. Frankfurter) and mouse anti-BrdU-FITC (1:20; Boehringer Mannheim). The secondary antibodies used in triple-labelling were donkey anti-species Texas red, Cy3 or Cy5 (1:250; Jackson Immunoresearch). Sections were imaged using a Bio-Rad MRC1000 confocal microscope.

Quantification. To examine the distribution of grafted cells, semiserial sections containing the OB, HC or CB (a 1 in 6 horizontal or coronal series) were stained immunohistochemically for BrdU and the number of BrdU-immunoreactive cells was quantified using stereological procedures¹¹. Quantitative data were compiled and analysed statistically. Error was expressed as the standard error of the mean (s.e.m.). The vector distance (straight-line distance) between the graft centre and the most distant AHP was measured (in micrometres) in the four anatomical axes from horizontal OB sections (five grafts, three sections per graft, one measurement per axis per graft), coronal hippocampal sections (five grafts,

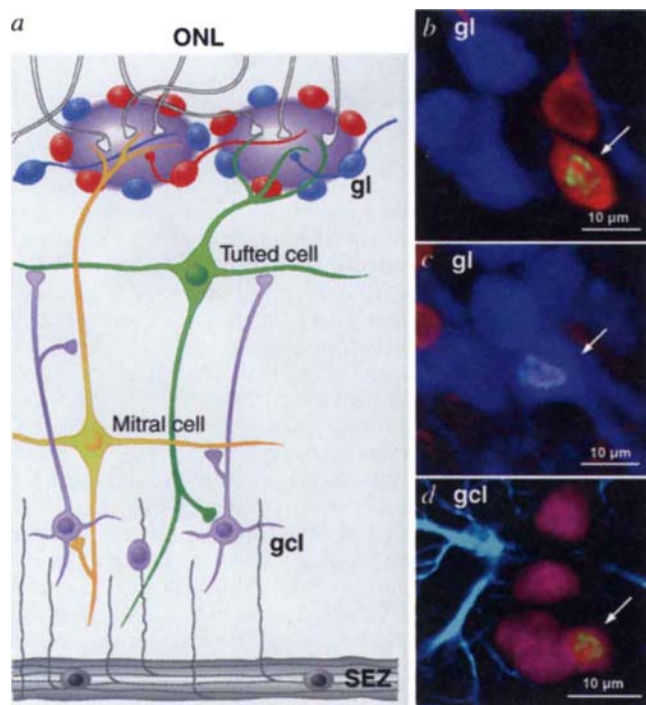


FIG. 3 Differentiation of grafted AHPs in OB. Schematic of adult OB (modified from ref. 30). ONL, olfactory nerve layer; gl, glomerular cell layer; gcl, granule cell layer; SEZ, subependymal zone (a). Neuronal phenotype of AHPs at eight weeks (b-d). Calbindin- (red) and TH- (blue) immunoreactive periglomerular neurons in the gl of OB (b, c). BrdU-immunoreactive AHPs express calbindin (yellow; arrow in b) and TH (light blue; arrow in c), a non-HC phenotype. NeuN-immunoreactive neurons (purple) forming the gcl of OB (d). Yellow indicates a grafted cell (arrow) in the gcl, double-labelled for NeuN and BrdU; turquoise indicates GFAP immunoreactivity.

three sections per graft, one medial/lateral measurement per graft) and cerebellar sections (five grafts, one section per graft, six medial/lateral measurements per graft). For coronal sections, rostrocaudal migration was calculated from the distance between sections.

Received 1 April; accepted 2 September 1996.

- Gage, F. H., Ray, J. & Fisher, L. J. *Annu. Rev. Neurosci.* **18**, 159–192 (1995).
- Kaplan, M. S. & Hinds, J. W. *Science* **197**, 1092–1094 (1977).
- Levison, S. V. & Goldman, J. E. *Neuron* **10**, 201–212 (1993).
- Luskin, M. B. *Neuron* **11**, 173–189 (1993).
- Altman, J. & Das, G. D. *J. Comp. Neurol.* **124**, 319–336 (1965).
- Bayer, S. A. *Exp. Brain Res.* **46**, 315–323 (1982).
- Schwartz-Levey, M., Chikaraishi, D. M. & Kauer, J. S. *J. Neurosci.* **11**, 3556–3564 (1991).
- Reynolds, B. A. & Weiss, S. *Science* **255**, 1707–1710 (1992).
- Lois, C. & Alvarez-Buylla, A. *Proc. Natl Acad. Sci. USA* **90**, 2074–2077 (1993).
- Richards, L. J., Kilpatrick, T. J. & Bartlett, P. F. *Proc. Natl Acad. Sci. USA* **89**, 8591–8595 (1992).
- Gage, F. H. *et al.* *Proc. Natl Acad. Sci. USA* **92**, 11879–11883 (1995).
- Palmer, T. D., Ray, J. & Gage, F. H. *Mol. Cell. Neurosci.* **6**, 474–486 (1995).
- Menezes, J. R., Smith, C. M., Nelson, K. C. & Luskin, M. B. *Mol. Cell. Neurosci.* **6**, 496–508 (1995).
- Lois, C. & Alvarez-Buylla, A. *Science* **264**, 1145–1148 (1994).
- Lois, C., Garcia-Verdugo, J.-M. & Alvarez-Buylla, A. *Science* **271**, 978–981 (1996).
- Rousselot, P., Lois, C. & Alvarez-Buylla, A. *J. Comp. Neurol.* **351**, 51–61 (1995).
- Hu, H., Tomasiewicz, H., Magnuson, T. & Rutishauser, U. *Neuron* **16**, 753–743 (1996).
- O'Rourke, N. A. *Neuron* **16**, 1061–1064 (1996).
- Edelman, G. M. & Rutishauser, U. *J. Supramol. Struct. Cell. Biochem.* **16**, 259–268 (1981).
- Miragall, F., Kadmon, G., Husman, M. & Schachner, M. *Dev. Biol.* **129**, 516–531 (1988).
- Seki, T. & Arai, Y. *Neurosci. Res.* **12**, 503–513 (1991).
- Gao, W. Q. & Hatten, M. E. *Science* **260**, 367–369 (1993).
- Vicario-Abejon, C., Cunningham, M. G. & McKay, R. D. G. *J. Neurosci.* **15**, 6351–6363 (1995).
- Mullen, R. J., Buck, C. R. & Smith, A. M. *Development* **116**, 201–211 (1992).
- Celio, M. R. *Neuroscience* **35**, 375–475 (1990).
- Halasz, N. *et al.* *Brain Res.* **126**, 455–474 (1977).
- Campbell, K., Olsson, M. & Björklund, A. *Neuron* **15**, 1259–1273 (1995).
- Fishell, G. *Development* **121**, 803–812 (1995).
- Rougon, G., Dubois, C., Buckley, N., Magnani, J. L. & Zollinger, W. J. *Cell Biol.* **103**, 2429–2437 (1986).
- Keveke, E. B. *Curr. Opin. Neurobiol.* **5**, 482–488 (1995).

ACKNOWLEDGEMENTS. We thank P. Eriksson, L. J. Fisher, M. Gage, M. B. Luskin, T. D. Palmer and H. K. Raymon for helpful criticism of the manuscript and S. Forbes for his help with animal care. This work was supported by separate grants from the NIA, NINDS, American Paralysis Association, Hollfelder Foundation, International Spinal Research Trust and the MRC of the Academy of Finland.

CORRESPONDENCE and requests for materials should be addressed to F.H.G. (e-mail: fgage@salik.edu).

Uncoupling cadherin-based adhesion from *wingless* signalling in *Drosophila*

Bénédicte Sanson, Phoebe White & Jean-Paul Vincent

Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

THE *Wnt* genes encode secreted glycoproteins used in intercellular communication at multiple steps during development. Signalling by *Wingless*, the *Drosophila Wnt-1* homologue, requires the activity of *Armadillo*^{1,2}, the homologue of vertebrate β -catenin³, which is a component of the cadherin/catenin complex at adherens junctions^{4,5}. The genetic link between *wingless* and *armadillo* suggests that cell fate specification and cell–cell adhesion might be controlled concurrently. For instance, in one extreme view, *Wingless* could specify cell fate entirely by modulating cell adhesion. Alternatively, it might signal independently of adherens junctions. To distinguish between these alternatives, we have expressed two polypeptides that have opposite effects on cadherin-dependent adhesion: full-length *Drosophila* E-cadherin and a dominant-negative truncated form. We found that overexpression of either construct mimics *wingless* phenotypes, thereby uncoupling changes in adhesion from signalling effects. We demonstrate that both constructs titrate *Armadillo* from a 'signalling' pool which is functionally distinct from the junctional pool.

Cadherins are Ca^{2+} -dependent adhesion molecules which, in

association with α - and β -catenin, constitute the major component of adherens junctions in vertebrates. β -catenin binds to both α -catenin and the intracellular domain of cadherin⁶, thereby providing a link between adhesion plaques and the cytoskeleton⁷, a link that is essential for cell adhesion^{8,9}. As expected from its homology with β -catenin, *Armadillo* interacts with *Drosophila* α -catenin and E-cadherin^{10,11} and is also required for cell adhesion¹². Experiments in tissue culture have shown that overexpression of native full-length cadherins (CAD) increases adhesion as expected¹³, whereas overexpression of the intracellular domain (CADⁱ) in *Xenopus* embryos decreases adhesion, indicating that CADⁱ has a dominant-negative effect¹⁴. These two polypeptides provide an opportunity to alter cadherin-dependent adhesion in opposite ways and see how *Wingless* signalling might be affected. If cadherin-mediated adhesion is a necessary part of *Wingless* signal transduction, overexpression of the two constructs should have opposite effects on *Wingless* signalling.

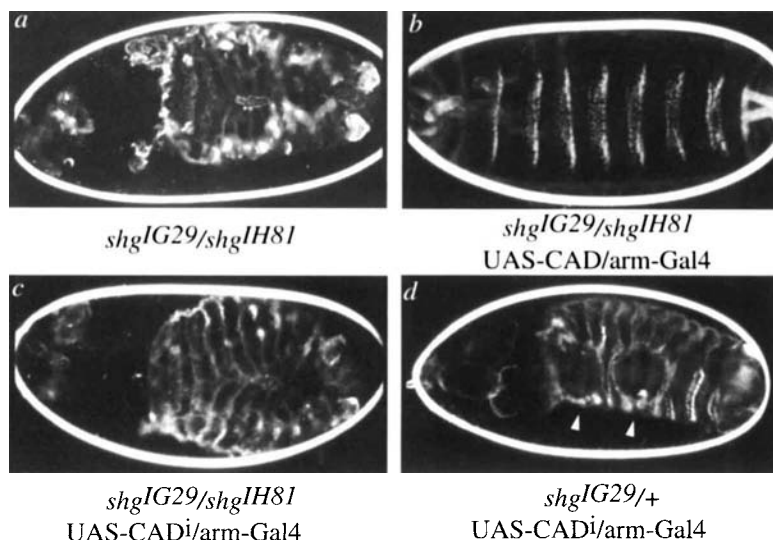
We used the Gal4/UAS system¹⁵ to overexpress *Drosophila* CAD and CADⁱ and compared the resulting phenotypes to known effects of *wingless* mutations. We first verified that CAD and CADⁱ do have opposite effects on cadherin function. First, we found that uniform expression of CAD (but not CADⁱ) fully rescues the cuticular phenotype of a strong mutant of *shotgun*, the gene encoding *Drosophila* E-cadherin^{16,17} (Fig. 1a–c). In contrast, although *shotgun* heterozygotes are completely normal, they develop *shotgun*-like phenotypes when they uniformly express CADⁱ, verifying that this construct has a dominant-negative effect upon E-cadherin function (Fig. 1d).

Despite their opposite effects on cadherin-based adhesion, we found that CAD and CADⁱ affect *Wingless* signalling similarly: each mimics the phenotype of *wingless* loss-of-function mutations both in the embryo and in adult appendages (Figs 2, 3). In the embryo, uniform expression of CAD or CADⁱ leads to a cuticular phenotype typical of intermediate *wingless* mutants (Fig. 2c, d). Overall, CAD transgenes are more effective than CADⁱ transgenes and expression from two copies of the most active CAD insertions fully mimics the phenotype of a *wingless* null allele (Fig. 2e, f). We confirmed that this phenocopy is genuine by looking at a molecular marker of *Wingless* function. Maintenance of ectodermal *engrailed* expression beyond 5 hours of development requires *wingless* activity^{18,19}. In embryos carrying two CAD insertions expressed uniformly, there is an almost complete loss of *engrailed* expression, as in *wingless* mutants (Fig. 2g–i).

Expression of the same constructs at lower levels, using different Gal4 drivers, leads to *wingless*-like phenotypes in the adult (Fig. 3). For example, expression of either the full-length or the truncated construct under the control of *engrailed*-Gal4 causes partial loss of the wing margin in the posterior compartment (Fig. 3a–c); the extent of margin loss is correlated with the strength of the responder transgenic line. Again this phenotype is similar to that of *wingless* hypomorphic mutants^{20,21}. Moreover, uniform expression of UAS-CADⁱ in imaginal discs can cause wing loss and notum duplication (Fig. 3d), the phenotype observed in *wingless*ⁱ mutants²². Thus, both constructs reduce *Wingless* signalling in the embryo and imaginal discs, despite the fact that they have opposite effects on cadherin-mediated adhesion. This strongly suggests that cell fate specification by *Wingless* is not primarily effected by modulation of cell adhesion.

How do overexpressed full-length and truncated cadherins mimic *wingless* mutations? A clue comes from the fact that both constructs retain the intracellular domain which contains a binding site for *Armadillo*^{23,24}, suggesting that CAD and CADⁱ may titrate it. Indeed, it has been shown²⁵ that overexpression of full-length cadherin or depletion of β -catenin both lead to ventralization in *Xenopus* embryos, hinting that excess cadherin might decrease the pool of β -catenin. Our finding that overexpressed CADⁱ, as well as overexpressed CAD, phenocopies *wingless* (or *armadillo*) mutants supports this hypothesis. If CAD and CADⁱ do reduce the level of *Armadillo* available for signalling, we expect their effects to be very sensitive to *armadillo* gene dosage. Indeed,

FIG. 1 Full-length cadherin (CAD) and the intracellular domain of cadherin (CADⁱ) have opposite effects on cadherin function. The two polypeptides (Fig. 2a) were expressed in *shotgun* homozygous and heterozygous embryos and the resulting cuticular phenotypes analysed. As expression levels vary from line to line, we identify each with a superscripted number. Recombinant lines carry two numbers. *a*, Cuticle from a *shg^{G29/shg^{IH81}}* embryo. Anterior is to the left, as in all subsequent pictures of embryos. This combination of alleles gives a strong phenotype characterized by a lack of head structure, a loss of the ventral cuticle, and the appearance of holes in the lateral cuticle. In some embryos, holes in the dorsal cuticle also form. *b*, Rescue of the *shotgun* phenotype by expression of full-length cadherin under the control of uniformly expressed Gal4. Genotype is UAS-CAD⁶ *shg^{G29/arm-Gal4¹¹ shg^{IH81}}*. Most of the embryos are rescued to hatching and all show a wild-type cuticle. *c*, Lack of rescue by overexpressed intracellular domain (CADⁱ). Genotype is UAS-CAD⁴ *shg^{G29/arm-Gal4¹¹ shg^{IH81}}*. Phenotype is identical to that of *shg^{G29/shg^{IH81}}* embryos. The same result is obtained by overexpressing two doses of UAS-CADⁱ (UAS-CAD^{4,5}) (not shown). *d*, Expression of CADⁱ causes adhesion defects in *shotgun* heterozygotes. Genotype is UAS-CAD⁴/arm-Gal4¹¹ *shg^{G29}*. Whereas the heterozygotes *shg^{G29/+}* show no defects, about 10–20% of the embryos of this genotype display strong head defects and holes in the ventral cuticle (arrowheads), typical of



intermediate *shotgun* mutations¹⁷. Note that the UAS-CAD⁴/arm-Gal4¹¹ combination used in this experiment gives no embryonic lethality in a wild-type background.

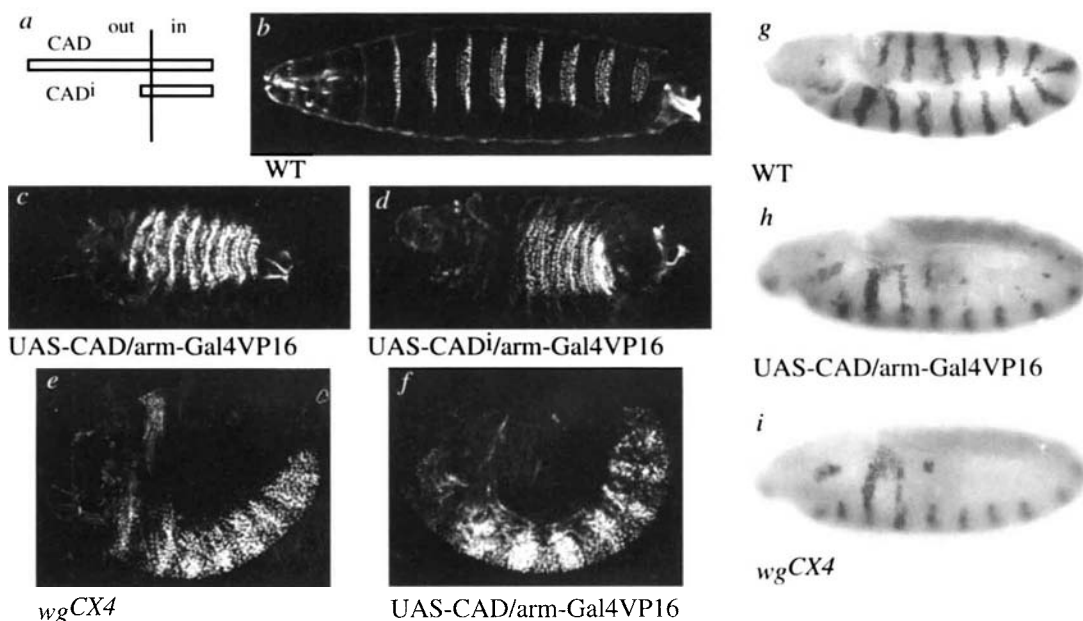


FIG. 2 Both full-length and truncated cadherin phenocopy *wingless* when overexpressed in embryos. *a*, Diagram of the two polypeptides. CADⁱ retains the transmembrane and the cytoplasmic domain. *b*, Wild-type cuticle. *c*, *d*, Uniform expression of either UAS-CAD or UAS-CADⁱ leads to segment polarity phenotypes typical of those seen in intermediate *wingless* mutants (see also below). Expression was activated by a Gal4-VP16 fusion protein expressed uniformly under the control of the *armadillo* promoter (arm-Gal4VP16). UAS-CAD^{4,5} and UAS-CAD⁹ inserts were used for this experiment. *e*, *f*, Two doses of UAS-CAD (*f*) (UAS-CAD^{5,9}) activated by arm-Gal4VP16 produce a cuticle phenotype identical to that of a *wingless* null mutant (*e*) (*wingless^{CX4}*). Wild-type embryos overexpressing CADⁱ (as in *d*) do not display holes in the ventral cuticle that are characteristic of *shg* mutants. However, these embryos do have *shg*-like defects in the head and dorsally. We do not yet understand why the ventral

region of these embryos appears to be insensitive to CADⁱ-induced adhesion defects. It is as if a ventral cuticle displaying a *wg*-like phenotype is somehow less likely to show adhesion defects. *g*, *h* and *i*, Two doses of UAS-CAD (UAS-CAD^{5,9}) activated by arm-Gal4VP16 cause loss of *engrailed* expression in a manner that resembles loss in *wingless* null mutant. In contrast to the situation in *wingless* null embryos (*i*), a few scattered cells maintain *engrailed* expression in embryos overexpressing CAD (*h*). These may arise because overexpression of CAD commences when *Wingless* may have already acted upon a small number of *engrailed*-expressing cells (*Wingless* begins to affect *engrailed* expression at early-stage 9 when arm-Gal4 begins to transactivate a test responder gene like UAS-*lacZ*; data not shown). *g*, Wild-type is shown for comparison. All embryos are at stage 11 and stained with anti-*Engrailed*.

FIG. 3 Overexpression of CAD or CADⁱ leads to *wingless*-like phenotypes in adults. *a*, Wild-type wing. Anterior and posterior are indicated. *b*, Wing from a fly of the genotype UAS-CAD²/*engrailed*-Gal4. *c*, Wing from a fly of the genotype UAS-CAD¹⁵/*engrailed*-Gal4. The phenotypes in *b* and *c* are both characterized by a loss of the wing margin in the posterior compartment where *engrailed*-Gal4 is expressed (A. Brand, personal communication). *d*, Scanning electron micrograph of the thoracic region of a fly expressing UAS-CAD^{4,5} under the control of *arm*-Gal4¹¹ (weak uniform expression in imaginal discs). With this combination, all the flies show defects in appendages; common defects include notum duplications or a drastic reduction of the wing down to a tiny winglet. Expression of the same recombined transgenes with an *arm*-Gal4VP16 driver is embryonic-lethal (Fig. 2*d*). Micrograph shows a supernumerary notum in the normal place of the wing (arrow).

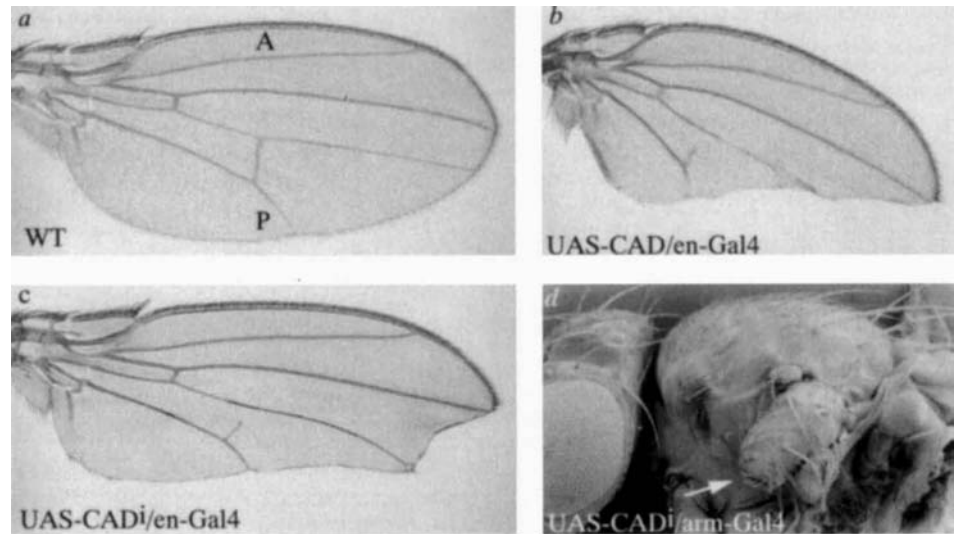
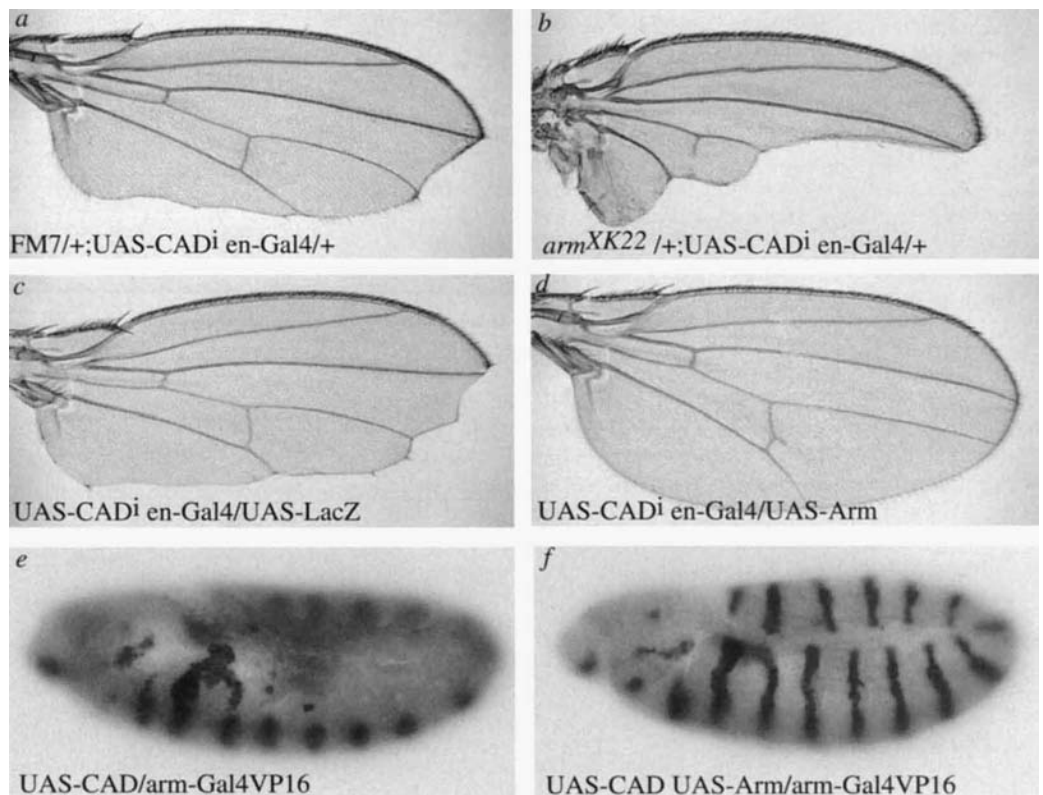


FIG. 4 CAD and CADⁱ phenotypes are enhanced by a decrease of *armadillo* function and rescued by exogenous Armadillo. *a*, Wing from a fly carrying a UAS-CAD¹⁰/*engrailed*-Gal4 recombinant chromosome. All the flies from this stock express CADⁱ in the posterior compartment and display this phenotype. A weak transgenic line was chosen to show clearly the enhancement obtained by halving *armadillo* gene dosage. *b*, Wing from a female heterozygous for a strong *armadillo* allele (*arm*^{XK22}/+) as well as carrying the same UAS-CAD¹⁰/*engrailed*-Gal4 recombinant chromosome. The wing phenotype of surviving flies is clearly enhanced (most die at pupal stages). Stronger UAS-CADⁱ transgenes are completely enhanced to lethality. *c*, Wing from a fly carrying a UAS-CAD¹⁵/*engrailed*-Gal4 recombinant chromosome. *d*, Wing from a fly carrying a UAS-Armadillo⁸ transgene in addition to the same recombinant chromosome as in *c*. The phenotype is rescued to wild-type. Rescue is not due to dilution of Gal4 activity because a control UAS-lacZ transgene has no rescuing activity (the wing shown in *c* actually carries UAS-lacZ). *e*, Loss of *engrailed* expression in an embryo expressing UAS-CAD⁵ under the control of *arm*-Gal4VP16. The phenotype arising from this UAS-CAD insert is variable, with some embryos continuing to express *engrailed* in a spotty fashion. *f*, The same



UAS-CAD transgene was recombined to UAS-Armadillo²³ and both transgenes were driven simultaneously by *arm*-Gal4VP16. Expression of *engrailed* is restored as no embryo from this cross showed abnormal *engrailed* expression.

we found that the dominant phenotypes are strongly enhanced in *armadillo* heterozygotes (Fig. 4a, b); in many cases, the enhancement causes lethality. For comparison, removing one dose of *wingless* does not enhance the effects of CAD or CADⁱ so dramatically (not shown). Conversely, increasing the amount of Armadillo protein rescues overexpression of CAD and CADⁱ. For example, the wing phenotypes are completely rescued when a UAS–*Armadillo* transgene is added (Fig. 4c–d). Moreover, in embryos, *engrailed* expression is maintained normally when UAS–*Armadillo* is expressed at the same time as UAS–CAD (Fig. 4e, f). We conclude that both constructs act by titrating Armadillo away from a signalling function in the Wingless pathway. This function must occur outside cadherin/catenin complexes; it could conceivably take place in the cytoplasm², elsewhere at the plasma membrane, or in the nucleus^{26,27}.

As we have shown, the level of cadherin clearly influences Wingless signalling. But as opposite changes in cadherin function affect Wingless signalling in the same direction, cadherin is probably not a primary component of the Wingless pathway. Thus we expect that Wingless signalling should not be eliminated by loss of cadherin activity. Indeed, we found *engrailed* expression to be normal in embryos mutant for *shotgun* (*shg*^{III}) and *shg*^{IG29}, results not shown). Moreover, *shotgun* mutant embryos display a cuticular phenotype which is distinct from that of *wingless* mutants^{16,17} (see also Fig. 1). However, these observations cannot exclude a potential role for cadherin in Wingless signalling as the zygotic mutants contain maternally encoded cadherin which may perdure during embryogenesis. Unfortunately, the maternal contribution cannot be removed because cadherin is necessary for oogenesis^{16,17}. As an alternative genetic test, we looked at *stardust* mutant embryos which have disrupted adherens junctions throughout the time when Wingless is required for maintenance of *engrailed* expression²⁸. Again, the maintenance of *engrailed* expression is normal (*sdt*^{EH}; data not shown), suggesting that Wingless signalling is unaffected by the disruption of zonulae adherentes. Thus, genetic analysis supports the hypothesis that cadherin function is not directly necessary for Wingless to signal.

Our results, considered with the recent demonstration that an adhesion-deficient Armadillo mutant can participate in signalling²⁷, indicate that Armadillo has a function in the Wingless pathway that is distinguishable from its role in adherens junctions (see also more indirect studies of *Xenopus* β -catenin^{25,26,29}). Nevertheless, the use of a common molecule in signalling and adhesion is probably significant. It is conceivable that, in addition to activating a signalling pathway, Wingless does modulate cell adhesion³⁰. Conversely, cadherin levels may influence the pool of Armadillo and hence Wingless signalling. We note here that in our rescue of *shotgun* mutants with CAD, the level of cadherin expression with the Gal4 system does not overshoot wild-type levels significantly because no rescued embryo showed a *wingless* phenotype (Fig. 1b); if overexpression were massive, a *wingless* phenotype would be expected whether or not the endogenous gene is present. Therefore, we suggest that Wingless signalling may be affected by endogenous levels of cadherin. Cadherin could possibly dampen signalling by Wingless and thus prevent spurious activation of the pathway. In other words, cadherin could buffer 'signalling Armadillo'. □

Methods

Gal4 driver lines. *arm*–Gal4 refers to a transgene expressing Gal4 under the control of the *armadillo* promoter, which is ubiquitous in embryos and imaginal discs. The *arm*–Gal4 transgene contains the *armadillo* promoter, the Gal4 cDNA taken from pGATb¹⁵ and the 3' untranslated region from the K10 gene. *Arm* > yellow > Gal4VP16 was a generous gift from L. Seugnet and M. Haenlin. This transgene can be converted to *arm*–Gal4VP16 by Flp, which excises the FRT cassette (> yellow >) inserted between the *armadillo* promoter and the Gal4–VP16 cDNA. As the 'flipped-out' stock is sick, we keep a fly stock with the cassette until use. Before use, the cassette is excised in the male germ line by introducing a testis-specific tubulin- β 2–Flp transgene (from K. Basler). These males are then crossed to homozygous UAS-responder females—about 45% of the progeny carry one copy of the flipped-out transgene (*arm*–Gal4VP16). The

engrailed–Gal4 flies were a generous gift from A. Brand.

Gal4-responsive lines. UAS–CAD was made by subcloning a DE-cadherin cDNA¹¹ as an EcoRI fragment into pUAST¹⁵. UAS–CADⁱ was made by assembling DNA encoding the signal peptide of *Drosophila* Argos, a Myc tag, and the C-terminal 208 amino acids of *Drosophila* E-cadherin, which include the transmembrane domain and the cytoplasmic domain. The construct was then subcloned as an EcoRI–NotI fragment into pUAST. UAS–*armadillo* was made by inserting an *armadillo* cDNA (E9) (from M. Peifer) into pUAST.

Staining and mounting of samples. Cuticles were mounted in Hoyer's and wings in Euparal medium. Anti-*engrailed* antibody (mAb4D9) was used according to standard procedures. Stained embryos were mounted in Permount.

Received 24 June; accepted 23 August 1996.

- Noordermeer, J., Klingensmith, J., Perrimon, N. & Nusse, N. *Nature* **367**, 80–83 (1994).
- Peifer, M., Sweeton, D., Casey, M. & Wieschaus, E. *Development* **120**, 369–380 (1994).
- Peifer, M. *J. Cell Sci.* **105**, 993–1000 (1993).
- Tsukita, S., Tsukita, S., Nagatuchi, A. & Yonemura, S. *Curr. Opin. Cell Biol.* **4**, 834–839 (1992).
- Stappert, J. & Kemler, R. *Curr. Opin. Neurobiol.* **3**, 60–66 (1993).
- Aberle, H. et al. *J. Cell Sci.* **3655**–3663 (1994).
- Rimm, D. L., Koslov, E. R., Kebraei, P., Cianci, C. D. & Morrow, J. S. *Proc. Natl Acad. Sci. USA* **92**, 8813–8817 (1995).
- Nagafuchi, A. & Takeichi, M. *Cell Regul.* **1**, 37–44 (1989).
- Ozawa, M., Baribault, H. & Kemler, R. *EMBO J.* **8**, 1711–1717 (1989).
- Oda, H. et al. *J. Cell Biol.* **121**, 1133–1140 (1993).
- Oda, H., Uemura, T., Harada, Y., Iwai, Y. & Tameichi, M. *Dev. Biol.* **165**, 716–726 (1994).
- Peifer, M., Orsulic, S., Sweeton, D. & Wieschaus, E. *Development* **118**, 1191–1207 (1993).
- Takeichi, M. *Annu. Rev. Biochem.* **59**, 237–252 (1990).
- Kintner, C. *Cell* **69**, 225–236 (1992).
- Brand, A. H. & Perrimon, N. *Development* **118**, 401–415 (1993).
- Uemura, T. et al. *Genes Dev.* **10**, 659–671 (1996).
- Tepass, U. et al. *Genes Dev.* **10**, 672–685 (1996).
- DiNardo, S., Sher, E., Heemskerk-Jongens, J., Kassis, J. A. & O'Farrell, P. H. *Nature* **332**, 604–609 (1988).
- Martinez Arias, A., Baker, N. & Ingham, P. W. *Development* **103**, 157–170 (1988).
- Couso, J. P., Bishop, S. A. & Martinez Arias, A. *Development* **120**, 621–636 (1994).
- Diaz Benjumea, F. & Cohen, S. M. *Development* **121**, 4215–4225 (1995).
- Sharma, R. P. & Chopra, V. L. *Dev. Biol.* **48**, 461–465 (1976).
- Ozawa, M., Ringwald, M. & Kemler, R. *Proc. Natl Acad. Sci. USA* **87**, 4246–4250 (1990).
- Aberle, H., Schwartz, H., Hoschuetzky, H. & Kemler, R. *J. Biol. Chem.* **271**, 1520–1526 (1996).
- Hearman, J. et al. *Cell* **79**, 791–803 (1994).
- Funayama, N., Fagotto, F., McCreary, P. & Gumbiner, B. M. *J. Cell Biol.* **128**, 959–968 (1995).
- Orsulic, S. & Peifer, M. *J. Cell Biol.* **134**, 1283–1300 (1996).
- Grawe, F., Wodarz, A., Lee, B., Knust, E. & Skaer, H. *Development* **122**, 951–959 (1996).
- Fagotto, F., Funayama, N., Glück, U. & Gumbiner, B. J. *Cell Biol.* **132**, 1105–1114 (1996).
- Hinck, L., Nathke, I. S., Papkoff, J. & Nelson, W. J. *Trends Biochem. Sci.* **19**, 538–542 (1994).

ACKNOWLEDGEMENTS. We thank J. Ahninger, M. Bienz, M. Freeman, R. Kay, P. Lawrence and D. St Johnston for discussions and critical reading of the manuscript, and L. Seugnet, M. Haenlin and A. Brand for unpublished fly stocks. B.S. is the recipient of an EMBO postdoctoral fellowship.

CORRESPONDENCE and requests for materials should be addressed to J.P.V. (e-mail: jpv@mrc-lmb.cam.ac.uk).

Nuclear import of the homeodomain protein Extradenticle in response to Wg and Dpp signalling

Richard S. Mann* & Muna Abu-Shaar†

* Department of Biochemistry and Molecular Biophysics, † Integrated Program in Cellular, Molecular and Biophysical Studies, Columbia University College of Physicians and Surgeons, 630 West 168th Street, New York, New York 10032, USA

IN *Drosophila*, Decapentaplegic (Dpp)¹ and Wingless (Wg)² are two secreted signalling proteins of the transforming growth factor (TGF)- β and Wnt families, respectively. Although both are often required during development, only a few downstream components of these signalling pathways have been described. Here we present evidence that in the embryonic midgut both signalling pathways control the subcellular localization of the homeodomain protein encoded by the *extradenticle* (*exd*) gene. Exd protein is predominantly nuclear in endoderm cells close to Dpp- and Wg-secreting cells of the visceral mesoderm, but is in the cytoplasm in more distant endoderm cells. Both *dpp* and *wg* are required for the nuclear localization of Exd in the endoderm,

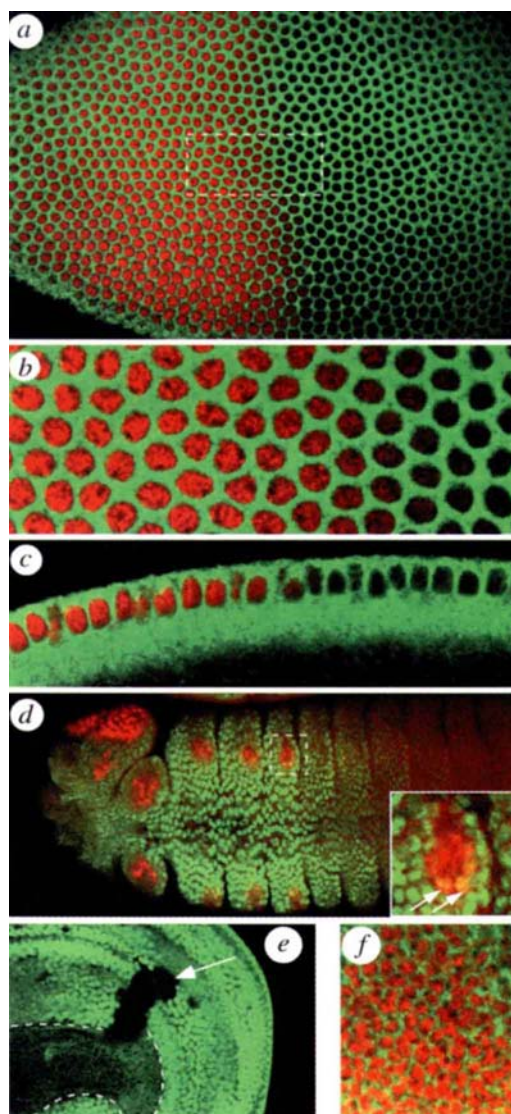


FIG. 1 Exd protein moves from the cytoplasm to the nucleus during embryogenesis. (Anterior is to the left in all figures). *a–c*, Wild-type blastoderm (stage 4) embryos stained for Exd (green) and Hunchback (Hb) (red) proteins; boxed region in *a* is enlarged in *b*. *c*, Optical cross-section of an embryo of similar age. Maternal Hb is present in nuclei in the anterior halves of blastoderm embryos²⁹, whereas Exd is in the cytoplasm throughout these embryos. *d*, Wild-type stage-13 embryo doubly stained for Exd (green) and Dll (red); the boxed region is enlarged in the inset. In all three thoracic segments, Dll is expressed in one cluster of cells per hemisegment²⁶. At this stage, Exd is absent from most of these nuclei; however, a subset of nuclei at the periphery of these clusters also contains Exd and thus appears yellow (arrows). *e*, Portion of a wing imaginal disc that contains a clone of homozygous *exd*^{P11} cells (arrow); this clone is adjacent to wing blade cells that have cytoplasmic Exd (outlined by the dashed line) and more peripheral cells that have nuclear Exd. *f*, Small portion of the wing blade region from a wing imaginal disc doubly stained for Dll (red) and Exd (green). Cytoplasmic Exd staining can be seen surrounding nuclear Dll staining.

whereas ectopic expression of *dpp* and *wg* expands the domain of nuclear Exd. Furthermore, the nuclear import of Exd correlates with the transcription of an *exd*-dependent reporter gene in the endoderm. Thus one mechanism by which extracellular signals might control pattern is by directing the graded nuclear localization of homeodomain proteins such as Exd that directly control the expression of target genes.

In young embryos, the protein Extradenticle (Exd) is principally derived from maternal transcripts³ and is uniformly distributed until gastrulation. Surprisingly, although Exd is a homeodomain

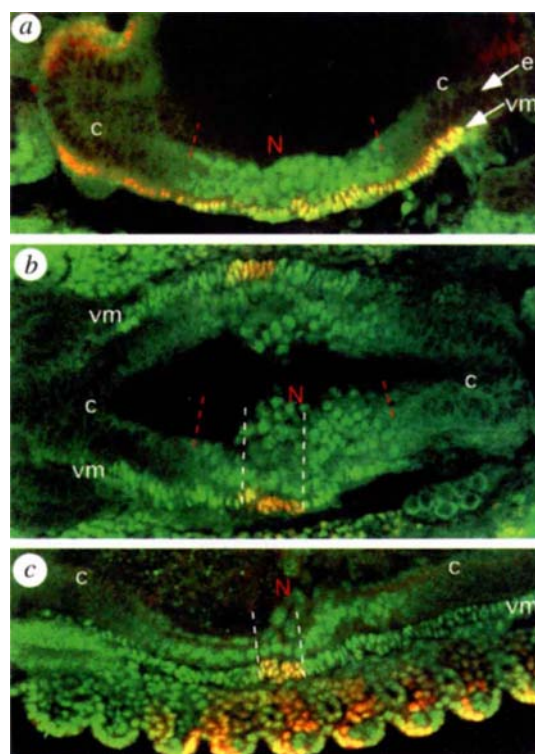


FIG. 2 The subcellular localization of Exd in the embryonic midgut. The visceral mesoderm (VM) is a thin layer of cells surrounding the gut endodermal cells (*e*), which are polyploid and therefore have large nuclei. *a*, Wild-type midgut of a stage-14 embryo carrying an enhancer trap in the *disconnected (disco)* gene stained for nuclear β -galactosidase (β -gal; red) and Exd (green). VM nuclei in the plane of focus appear yellow because of overlapping expression. In the endoderm, nuclear Exd (N) is visible only in the centre; the anterior and posterior portions of the endoderm have cytoplasmic Exd (*c*). Red dashed lines delineate the portion of the endoderm where Exd is observed in nuclei; 'N' marks the position of the most strongly staining nuclei. In wild-type stage-12 embryos, before fusion of the posterior and anterior midgut primordia, a subset of cells in both primordia have Exd localized in nuclei (data not shown). As with *labial* expression at this stage³⁰, this nuclear Exd, some of which may persist after fusion, is independent of Dpp or Wg signalling. *b*, Wild-type midgut of a stage-14 embryo stained for Ubx (red) and Exd (green). PS 7 VM nuclei, indicated by white dashed lines, appear yellow owing to overlapping expression. *c*, Midgut from an embryo that is homozygous for the null mutation *exd*^{P11} and stained for Ubx (red) and Exd (green).

protein that binds DNA and functions as a transcription factor⁴, it seems to be excluded from nuclei in blastoderm embryos (Fig. 1*a–c*). Following gastrulation, Exd is imported into specific subsets of nuclei (data not shown) and by mid-embryogenesis (stage 13), Exd is nuclear in most ventral ectodermal cells of the thorax and head (Fig. 1*d*). In the abdominal segments, Exd is observed at lower levels, consistent with the distribution of *exd* transcripts at this stage³ (Fig. 1*d*). Also at this stage, Exd is apparently absent from the nuclei of most cells that express the homeodomain protein encoded by the *Distal-less (Dll)* gene⁵ (Fig. 1*d*). This partially reciprocal relationship between Exd and Dll is maintained in wing and leg imaginal discs (Fig. 1*f*, and data not shown). Cytoplasmic and nuclear staining using the anti-Exd antibody must faithfully reflect the distribution of Exd protein because no staining was detectable in cells homozygous for an *exd* null allele (Fig. 1*e*). Regulation of the subcellular localization of Exd may explain why ubiquitous *exd* expression does not result in morphological defects⁶ despite the non-uniform distribution of nuclear Exd protein⁷.

To investigate the significance of the cytoplasmic versus nuclear localization of Exd, we examined its distribution in the embryonic midgut, where a genetic requirement for *exd* function has been

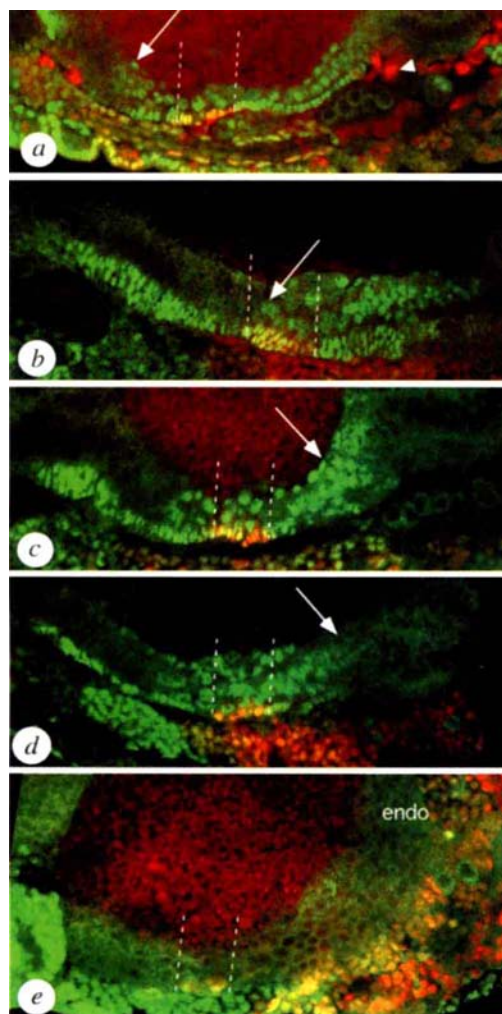


FIG. 3 Exd nuclear localization requires *dpp* and *wg*. All panels are confocal images of stage 13 or 14 midguts stained for Ubx (red) and Exd (green). The extent of Ubx expression in the VM (yellow), which marks PS 7, is indicated by dashed lines. a and b, Sibling embryos, stained simultaneously, which are either wild-type (a; arrowhead points to β -gal staining expressed from a balancer chromosome) or homozygous for the *dpp*²⁴ mutation (b). Arrows indicate the anterior-most nuclei in the endoderm that clearly stain positively for Exd. c and d, Sibling embryos, stained simultaneously, which are either wild-type (c) or homozygous for the *wg*^{cx4} mutation (d). Arrows indicate approximately equivalent positions in the endoderm. e, Midgut from an embryo doubly mutant for *dpp* and *wg*. In these embryos Exd is predominantly in the cytoplasm throughout the endoderm (endo).

described⁸⁻¹⁰. The midgut is derived from two germ layers: an inner endodermal layer of large polyploid cells, and an outer visceral mesoderm layer of smaller diploid cells (VM)^{11,12}. To identify the VM, we stained embryos containing an enhancer-trap (expressing nuclear β -galactosidase (β -gal)) in the *disconnected* (*disco*) gene, which is expressed in all VM cells¹³. In stage-14 embryos, Exd was seen in the nuclei of most VM cells (Fig. 2a). In contrast, in the adjacent endoderm, Exd was nuclear only in centrally positioned cells along the anterior-posterior axis (Fig. 2a-c). Anterior and posterior to these cells, the subcellular distribution of Exd gradually shifted from the nucleus to the cytoplasm until, in cells at the two ends of the midgut endoderm, Exd appeared to be entirely in the cytoplasm and absent from nuclei. To determine whether the different levels of nuclear Exd were the result of a post-transcriptional event, we stained embryos that were homozygous for an *exd*-null mutation, in which all Exd protein must be derived from maternal transcripts. Although

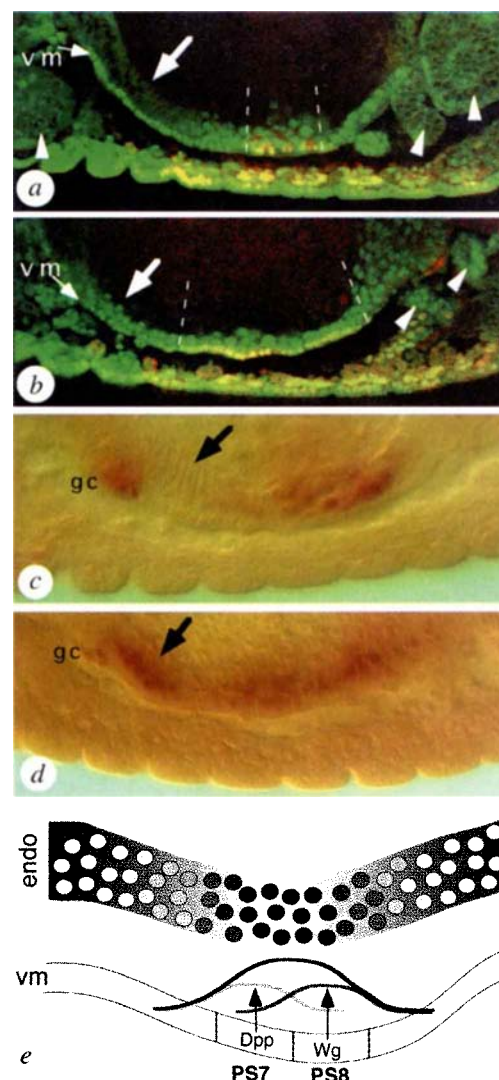


FIG. 4 Ectopic Dpp can drive Exd into nuclei. a and b, Sibling embryos, stained simultaneously for Ubx (red) and Exd (green), which expressed Dpp in a wild-type pattern (a) or ubiquitously (b). c and d, Sibling embryos containing the *3Xrpt3-lacZ* reporter gene, stained simultaneously for β -gal, which expressed Dpp in a wild-type pattern (c) or ubiquitously (d). In embryos that have a wild-type *dpp* expression pattern, anterior endoderm cells have Exd predominantly in the cytoplasm (large arrow in a) and no detectable expression of the reporter gene (large arrow in c). In addition, several other tissues visible in a appear to have predominantly cytoplasmic staining (arrowheads). In embryos that have ubiquitous *dpp* expression, anterior endoderm cells have Exd predominantly in nuclei (large arrow in b) and express the reporter gene (large arrow in d). In addition, other cells visible in b have nuclear Exd (arrowheads). Ectopic *dpp* also induces Ubx expression in the VM²² (seen as an increase in yellow VM nuclei) and allows unambiguous identification of these embryos. The *3Xrpt3-lacZ* reporter gene also expresses *lacZ* in the gastric caeca primordia (gc) (PS 3 of the VM)¹⁰; this expression, visible in c and d, serves as a marker for the anterior midgut. In summary (e), these data indicate that a combined gradient of Wg and Dpp (black curve), secreted from the VM, generates a gradient of nuclear Exd in the midgut endoderm (endo). Cytoplasmic Exd is depicted as the light-to-dark-grey background; nuclear Exd is shown as the grey-to-black circles.

these embryos have lower amounts of Exd, we observed a normal distribution of nuclear and cytoplasmic Exd, demonstrating that variations in the levels of nuclear Exd could not be due to differences in transcription (Fig. 2c, and data not shown).

The peak of nuclear Exd in the endoderm appeared to be in cells closest to the boundary between parasegment (PS) 7 and PS 8 of the VM (Fig. 2b, c). PS 7 VM cells, which were identified by

staining for the Hox protein Ultrabithorax (Ubx), secrete the TGF- β protein Dpp and PS 8 VM cells secrete Wg¹⁴⁻¹⁶. Dpp secreted from the VM is necessary for expression of the Hox gene *labial* in adjacent endoderm cells¹⁶⁻¹⁸. Curiously, in stage-11 embryos the protein Labial is predominantly in the cytoplasm in endodermal precursor cells of the posterior midgut primordia¹⁷. In older (stage-13) embryos, Labial is nuclear and, within its expression domain, is in a posterior-to-anterior gradient that parallels the anterior half of the Exd nuclear gradient¹⁷. The increased Labial in the posterior portion of its expression domain requires Wg¹⁷, and different concentrations of Wg may generate cell-type diversity in the endoderm¹⁹. Thus a combination of at least two signalling molecules, Dpp and Wg, that are secreted from the mesoderm is required to pattern the underlying endoderm.

We therefore tested whether one of the downstream events mediated by Dpp or Wg was to induce nuclear import of Exd in the endoderm, using Ubx expression in the VM to mark the position of PS 7. In embryos homozygous for the *dpp*⁵⁴ mutation, in which *dpp* function is specifically reduced in the midgut^{17,18,20}, the anterior extent of nuclear Exd in the endoderm was decreased (Fig. 3a, b). In contrast, Exd remained nuclear in endoderm cells that were close to PS 8 of the VM. In the midguts of embryos homozygous for a *wg*-null allele, *wg*^{CX4}, the amount of nuclear localized Exd in the endoderm was also decreased, especially in posterior cells (Fig. 3c, d). In these embryos, peak nuclear Exd levels were found in cells adjacent to PS 7 VM cells, which still secrete Dpp. In stage-14 embryos that were doubly mutant for both *dpp* and *wg*, Exd appeared to be predominantly in the cytoplasm throughout the endoderm (Fig. 3e). Thus both Wg and Dpp are required for the normal subcellular distribution of Exd in the midgut endoderm.

A second test of whether *dpp* and *wg* control nuclear import of Exd is to express these signalling molecules ectopically. When *dpp* was ubiquitously expressed using the GAL4 system²¹, nuclear Exd was observed throughout the endoderm (Fig. 4a, b). Ectopic expression of Wg also induced nuclear import of Exd, but less extensively than did Dpp (data not shown). Because ectopic *dpp* expression also induces *wg* expression²², these data, together with the loss-of-function results, suggest that a combination of Dpp and Wg signals is, at least in part, responsible for inducing the cytoplasmic-to-nuclear translocation of Exd (Fig. 4e).

To determine whether control of the subcellular localization of Exd regulates gene expression, we characterized the effect of ectopic Dpp on the activity of a reporter gene, *3Xrpt3-lacZ*, which is normally expressed in the same endoderm cells as *labial*¹⁰. The enhancer present in *3Xrpt3-lacZ* cooperatively binds Labial and Exd *in vitro* and requires *labial* and *exd* functions *in vivo*^{10,23}. Ubiquitous *dpp* expression induced transcription of this reporter gene throughout the endoderm (Fig. 4c, d), contrasting with the inability of ubiquitous *labial* expression to turn on this reporter gene in additional endoderm cells¹⁰. These results suggest that ability of Labial to induce transcription of *3Xrpt3-lacZ* and probably other target genes in the endoderm is limited by the presence of nuclear Exd.

To our knowledge, these results provide the first example of a homeodomain protein whose activity, like that of the *rel* transcription factors NF- κ B and Dorsal²⁴, is controlled by its subcellular localization. We suggest that in the embryonic midgut the nuclear import of Exd is an important downstream component of both the Dpp and Wg signals. Both of these signalling molecules appear to generate a gradient of nuclear-localized Exd (Fig. 4e). Although our experiments focused on the embryonic midgut, Exd's subcellular localization is apparently regulated in many tissues during development (Fig. 1, and data not shown). However, the nuclear localization of Exd does not always correlate with Wg or Dpp signals, suggesting that other signals control the subcellular localization of Exd elsewhere in development. Thus the cellular context is probably critical for determining the response to a particular signal. Likewise, a given signal often has different potentials to activate NF- κ B, depending on the cell type²⁵.

Because Exd and the closely related mammalian Pbx proteins

can increase the target-gene specificity of Hox homeoproteins through cooperative DNA binding⁴, we propose that regulation of Exd's subcellular localization is an important signal-dependent mechanism for controlling Hox and Exd target-gene expression. Generating gradients of nuclear transcription factors may be a mechanism frequently used by secreted signalling molecules to generate pattern in a responding tissue. □

Methods

Immunofluorescence. An N-terminal, histidine-tagged 143-amino-acid fragment of *exd* was overexpressed in *Escherichia coli* using the pET14b vector and purified by standard procedures¹⁰. A polyclonal rabbit antibody directed against this protein was prepared by Cocalico, Inc. (Pennsylvania). A 1:10 dilution of the rabbit serum was preabsorbed with an equal volume of fixed wild-type *Drosophila* embryos and then further diluted by 1:100 for all experiments. The secondary antibody used to detect the Exd antibody for most experiments was an FITC-conjugated goat anti-rabbit polyclonal antibody (Jackson). The anti-Dll (from S. Cohen)²⁶ and anti-Hb (from G. Struhl) antibodies were polyclonal mouse and rat antibodies, respectively; the secondary antibodies were conjugated with Texas red (Jackson). Immunofluorescence was done using standard procedures. All photomicrographs (except those shown in Fig. 4c, d) are confocal images. Imaginal discs were dissected from third-instar larvae and prepared for immunohistochemistry using standard procedures.

Fly stocks. *exd*^{XP11} is a null allele that has a stop codon after codon 61 (ref. 6). Clones of *exd*^{XP11} cells were generated by the FLP recombination system²⁷. *exd*^{XP11} embryos (Fig. 2) were identified by the absence β -gal expression from an *eve-lacZ* reporter gene on an FM7 balancer. *dpp*⁵⁴ embryos were identified by the absence of an *actin-lacZ* reporter gene on a CyO balancer, and *wg*^{CX4} embryos were identified by their morphology. Whenever possible, sibling embryos at similar stages of embryogenesis were stained in the same tube and imaged for about the same time. *wg dpp* double-mutant embryos were homozygous for the two null mutations *dpp*^{h61} and *wg*^{CX4}; the early requirement for *dpp* function for dorsal-ventral patterning was provided by a *dpp*^{hmn} minigene on the X chromosome, which does not provide any other *dpp* functions²⁸. The *3Xrpt3-lacZ* reporter gene behaves as a *labial* autoregulatory element *in vivo* and contains three tandem binding sites for a Labial-Exd heterodimer; both the Exd and Labial binding sites are required for *in vivo* function^{10,23}. Ectopic *dpp* expression was generated using the GAL4 system²¹. GAL4-UAS-*dpp* flies (stock 42B.4; Bloomington Stock Center) and two GAL4 driver lines were used: a ubiquitous GAL4 line, MS444 (from G. Morata), and the 24B line²¹, which expresses GAL4 in the mesoderm (not shown). Both driver lines produced similar results in the midgut. The driver lines were crossed to GAL4-UAS-*dpp* flies and F₁ embryos stained for Exd and Ubx (with mAbFP3.38) or β -gal (a rabbit anti- β -gal antibody; Cappel).

Received 24 June; accepted 5 September 1996.

1. Gelbart, W. M. *Development* **107**, 65–74 (1989).
2. Klingensmith, J. & Nusse, R. *Dev. Biol.* **166**, 396–414 (1994).
3. Rauskolb, C., Peifer, M. & Wieschaus, E. *Cell* **74**, 1–20 (1993).
4. Mann, R. S. & Chan, S.-K. *Trends Genet.* **12**, 258–262 (1996).
5. Cohen, S. M., Bronner, G., Kuttner, F., Jurgens, G. & Jackle, H. *Nature* **338**, 432–434 (1989).
6. Rauskolb, C., Smith, K., Peifer, M. & Wieschaus, E. *Development* **121**, 3663–3673 (1995).
7. Gonzalez-Crespo, S. & Morata, G. *Development* **121**, 2117–2125 (1995).
8. Chan, S.-K., Jaffe, L., Capovilla, M., Botas, J. & Mann, R. S. *Cell* **78**, 603–615 (1994).
9. Rauskolb, C. & Wieschaus, E. *EMBO J.* **13**, 3561–3569 (1994).
10. Chan, S.-K., Pöpperl, H., Krumlauf, R. & Mann, R. S. *EMBO J.* **15**, 2477–2488 (1996).
11. Poulson, D. F. in *Biology of Drosophila* (ed. Demerec, M.) (Wiley, New York, 1950).
12. Tepass, U. & Hartenstein, V. *Development* **120**, 579–590 (1994).
13. Lee, K. J., Freeman, M. & Steller, H. *EMBO J.* **10**, 817–826 (1991).
14. Bienz, M. & Tremml, G. *Nature* **333**, 576–578 (1988).
15. Tremml, G. & Bienz, M. *EMBO J.* **8**, 2677–2685 (1989).
16. Reuter, R., Panganiban, G. E., Hoffmann, F. M. & Scott, M. P. *Development* **110**, 1031–1040 (1990).
17. Immergluck, K., Lawrence, P. A. & Bienz, M. *Cell* **62**, 261–268 (1990).
18. Panganiban, G. E., Reuter, R., Scott, M. P. & Hoffmann, F. M. *Development* **110**, 1041–1050 (1990).
19. Hoppler, S. & Bienz, M. *EMBO J.* **14**, 5016–5026 (1995).
20. Hursh, D. A., Padgett, R. W. & Gelbart, W. M. *Development* **117**, 1211–1222 (1993).
21. Brand, A. & Perrimon, N. *Development* **118**, 401–415 (1993).
22. Staehling-Hampton, K. & Hoffmann, F. M. *Dev. Biol.* **164**, 502–512 (1994).
23. Chan, S.-K. & Mann, R. S. *Proc. Natl Acad. Sci. USA* **93**, 5223–5228 (1996).
24. Rushlow, C. & Warrior, R. *Bioessays* **14**, 89–95 (1992).
25. Grilli, M., Chiu, J. J. & Lenardo, M. J. *Int. Rev. Cytol.* **143**, 1–62 (1993).
26. Vachon, G. et al. *Cell* **71**, 437–450 (1992).
27. Xu, T. & Rubin, G. *Development* **117**, 1223–1237 (1993).
28. Hoffmann, M. & Goodman, W. *Genes. Dev.* **1**, 615–625 (1987).
29. Tautz, D. *Nature* **332**, 281–284 (1988).
30. Diederich, R. J., Merrill, V. K., Pultz, M. A. & Kaufman, T. C. *Genes Dev.* **3**, 399–414 (1989).

ACKNOWLEDGEMENTS. We thank K. Basler, J. Jiang, K. Lee, G. Morata, G. Struhl and the Bloomington Stock Center for flies; S. Cohen and G. Struhl for antibodies; and I. Greenwald, T. Jessell, G. Struhl and A. Tomlinson for comments on this manuscript. This work was funded by grants from the NIH and Human Frontiers Science Program to R.S.M.

CORRESPONDENCE and requests for materials should be addressed to R.S.M. (e-mail: rsm10@columbia.edu).

Flux coupling in a neuronal glutamate transporter

Noa Zerangue & Michael P. Kavanaugh

Vollum Institute, Oregon Health Sciences University, Portland, Oregon 97201, USA

SYNAPTIC transmission is commonly terminated by diffusion and reuptake of neurotransmitter from the synaptic cleft. Glutamate reuptake prevents neurotoxicity and sets the lower limit for the concentration of extracellular glutamate, so it is important to understand the thermodynamics of this process. Here we use voltage clamping with a pH-sensitive fluorescent dye to monitor electrical currents and pH changes associated with flux of glutamate mediated by the human neuronal glutamate transporter EAAT3. In contrast to a previous model¹, we find that three sodium ions and one proton are cotransported with each glutamate ion into the cell, while one potassium ion is transported out of the cell. This coupling can support a transmembrane glutamate concentration gradient ($[Glu]_{in}/[Glu]_{out}$) exceeding 10^6 under equilibrium conditions, and would allow the transporter to continue removing glutamate over a wide range of ionic conditions.

Glutamate uptake is driven by cotransport of sodium and countertransport of potassium^{2,3}. In addition, movement of pH-changing ions occurs during transport^{1,4,5}. Translocation of net positive charge with glutamate allows measurement of transport currents under voltage-clamp conditions, but this current also includes a thermodynamically uncoupled chloride flux activated by glutamate⁶⁻¹⁰. This uncoupled conductance pathway is permeable to anions in addition to Cl^- , including Br^- , I^- , NO_3^- and SCN^- (refs 7, 9). A countertransported pH-changing anion like OH^- or HCO_3^- (ref. 1) might therefore permeate through this channel-like pathway, resulting in a pH change during glutamate transport. To distinguish this possibility from thermodynamically coupled flux of a pH-changing ion, pH changes were monitored in oocytes expressing the human neuronal transporter EAAT3

(EAAT3)^{11,12} during uptake of glutamate under voltage clamp (membrane voltage, $V_m = -20$ mV), in conditions in which the OH^- ion gradient was directed either outwardly ($pH_o = 7.5$) or inwardly ($pH_o = 8.5$). As reported previously, the glutamate transport current was decreased by raising the external pH¹³, but uptake still resulted in intracellular acidification even with an inwardly directed electrochemical gradient for OH^- ions (Fig. 1a; control uninjected oocytes, Fig. 1b). This result verifies that flux of a pH-changing ion is thermodynamically coupled to glutamate flux^{5,10}. To examine whether a cotransported proton or a countertransported pH-changing anion is involved in this process, intracellular pH changes during uptake of the acidic substrates L-glutamate and L-cysteate were compared with changes during uptake of L-cysteine, a structurally related substrate of the neuronal glutamate transporter with a higher pK value, which is transported as a neutral zwitterion¹⁴. Unlike L-cysteate and L-glutamate, transport of an equivalent amount of L-cysteine did not result in a marked intracellular acidification, which would be expected if OH^- were countertransported (Fig. 1c). Instead, it seems likely that a proton is cotransported with the amino acid as a thiolate, sulphate, or carboxylate ion pair, leading to an intracellular pH change dependent on the pK of the amino acid. After intracellular release, cysteine ($pK = 8.3$) would remain predominantly protonated, whereas glutamic or cysteic acid ($pK < 5$) would release this proton, assuming an intracellular pH of 7.3 (ref. 15).

Thermodynamic coupling of the electrochemical gradients for H^+ , Na^+ , and K^+ to that of glutamate was examined quantitatively by measuring the effects of these gradients on the transport current reversal potential. In chloride-free Ringer solution, superfusion of L-glutamate produced an inward current in voltage-clamped Cl^- -depleted oocytes expressing EAAT3 which was blocked by the competitive antagonist kainate, a non-transported structural analogue of glutamate (Fig. 2a). Intracellular excitatory amino-acid concentrations in *Xenopus* oocytes are roughly 12 mM (ref. 14), similar to values reported for neurons¹⁶. Reverse transport of glutamate and an outward current is induced in salamander retinal glia cells by extracellular K^+ superfusion in the absence of external glutamate¹⁷. Reverse transport by the human neuronal transporter is similarly induced by potassium and this efflux is

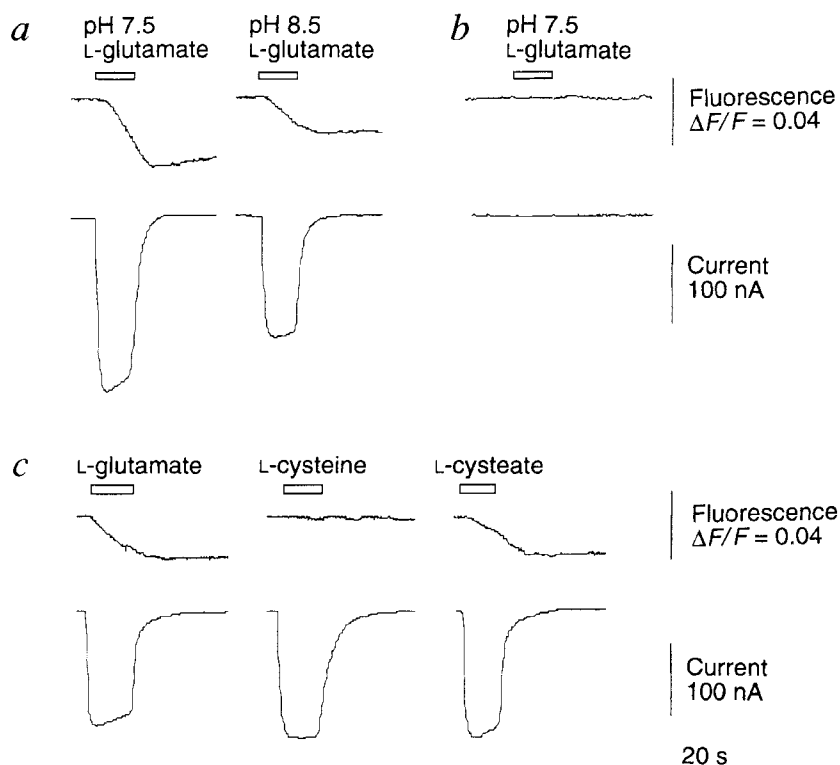


FIG. 1 *a*, A representative oocyte expressing EAAT3 was voltage clamped at -20 mV and superfused with Ringer solution at pH 7.5 (proton gradient inwards) or 8.5 (proton gradient outwards). Application of $300 \mu M$ L-glutamate induced an inward transport current and fluorescence decrease, reflecting intracellular acidification under each condition. *b*, No electrical current or pH change is induced by glutamate superfusion in control uninjected oocytes. *c*, Acidification associated with transport is dependent on the pK of the transported amino acid. Intracellular pH change in a representative oocyte expressing EAAT3 during sequential superfusion with Ringer solution (pH 7.5) containing L-glutamate ($300 \mu M$), L-cysteine ($1,000 \mu M$) and L-cysteate ($300 \mu M$). Similar results were seen in seven cells.

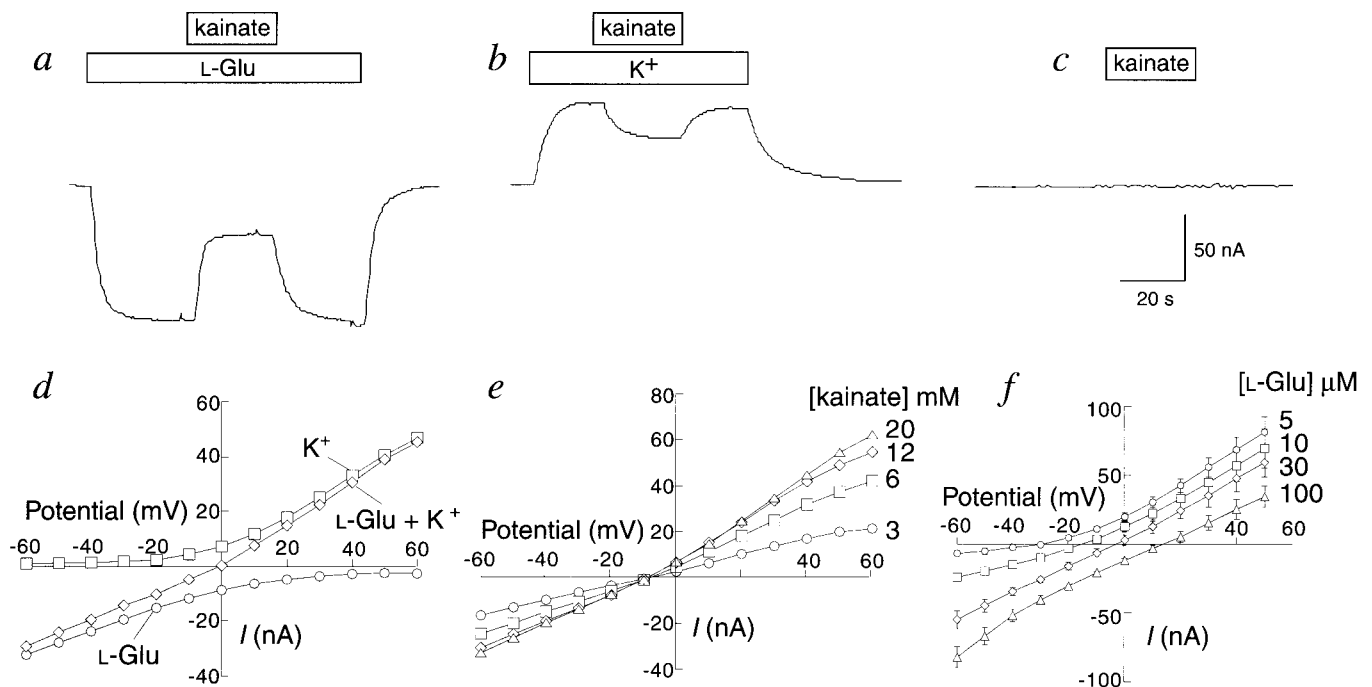


FIG. 2 *a*, EAAT3-mediated glutamate uptake ($10\ \mu\text{M}$ L-glutamate application; Cl^- and K^+ -free Ringer) is antagonized by co-applied kainate ($5\ \text{mM}$). *b*, Reverse transport current activated by addition of K^+ ($40\ \text{mM}$) without L-glutamate is similarly antagonized by kainate. *c*, Kainate has no effect on the membrane conductance in K^+ - and glutamate-free Ringer. Membrane potential, $-20\ \text{mV}$. *d*, Voltage-dependence of forward and reverse transport currents blocked by $5\ \text{mM}$ kainate in a representative cell. Efflux current (squares) recorded in $40\ \text{mM}$ K^+ in the absence of external glutamate, influx

current (circles) recorded in presence of $10\ \mu\text{M}$ glutamate in the absence of external K^+ . Transport current in conditions allowing bidirectional flux (diamonds) recorded in Ringer containing $10\ \mu\text{M}$ glutamate + $40\ \text{mM}$ K^+ . *e*, Reversal potential of transport current ($[\text{Glu}]_o = 10\ \mu\text{M}$, $[\text{K}^+]_o = 40\ \text{mM}$) was independent of fraction of the current blocked. *f*, Increasing $[\text{Glu}]_o$ results in a positive shift of the transport current reversal potential ($[\text{K}^+]_o = 40\ \text{mM}$, $[\text{Na}^+]_o = 60\ \text{mM}$). Points represent mean $\pm$ s.e.m. of 5–7 cells.

antagonized by kainate (Fig. 2*b*). Potassium and kainate exhibited the same actions on release of radioactivity from oocytes expressing EAAT3 which had been equilibrated with $[\text{H}^3]\text{p}$ -aspartate (data not shown). Kainate had no effect on the membrane conductance in oocytes expressing EAAT3 in the absence of external potassium or glutamate (Fig. 2*c*). The forward-transport current recorded in the presence of extracellular glutamate with no external potassium was inwardly rectified, as expected for unidirectional influx of glutamate (Fig. 2*d*). Conversely, the reverse-transport current induced by raising the external potassium concentration ($40\ \text{mM}$) in the absence of external glutamate was outwardly rectified (Fig. 2*d*). In the presence of external potassium and glutamate, a condition allowing either forward or backward transport, the transport current was inward at negative potentials and outward at positive potentials (Fig. 2*d*). The slope of the blocked transport current–voltage plot increased with increasing concentrations of kainate, whereas the zero-current potential remained constant, indicating that kainate selectively blocked transport without affecting the transport equilibrium (Fig. 2*e*).

We consider a thermodynamic reaction cycle involving coupled translocation of n_{Glu} glutamate ions, n_{Na} sodium ions, and n_{H} hydrogen ions from outside to inside, and n_{K} potassium ions from inside to outside. At equilibrium, the free energies of the coupled ions define a zero-flux equation relating the membrane potential Ψ to the transmembrane ion gradients:

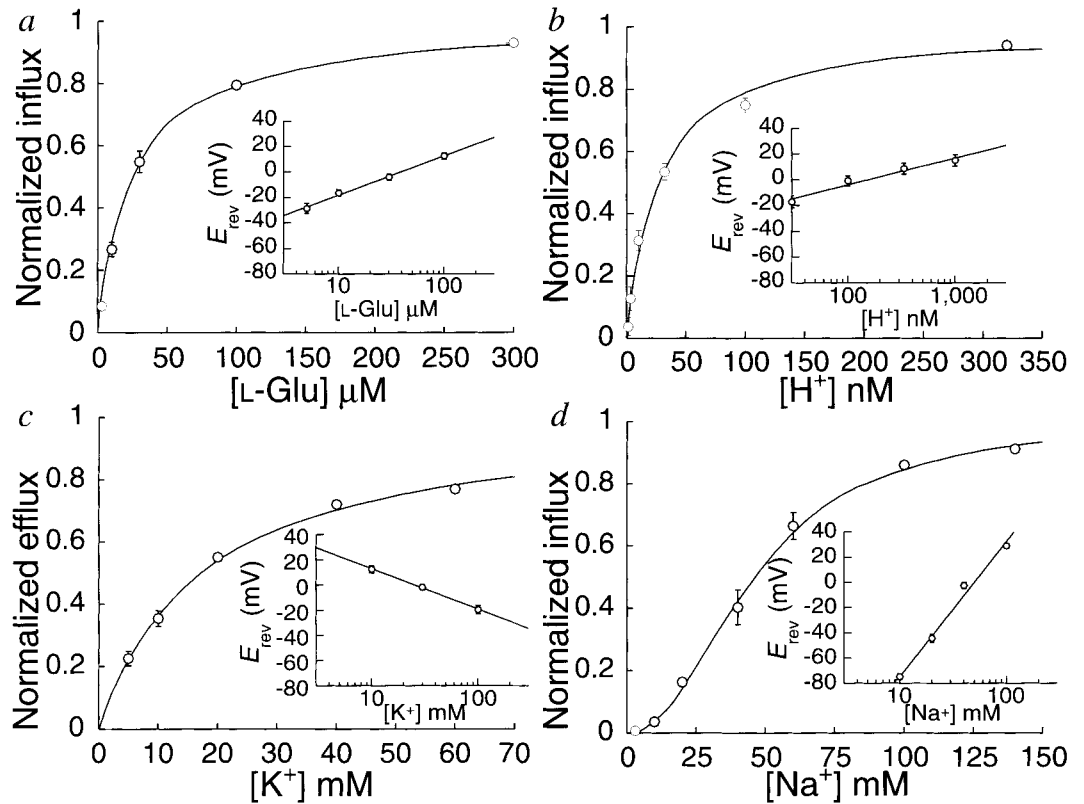
$$\Psi = (RT/F(n_{\text{Na}} + n_{\text{H}} - n_{\text{K}} - n_{\text{Glu}})) \ln([[\text{Na}]_o]/[\text{Na}]_i)^{n_{\text{Na}}} \times ([[\text{Glu}]_o]/[\text{Glu}]_i)^{n_{\text{Glu}}} ([[\text{H}]_o]/[\text{H}]_i)^{n_{\text{H}}} ([[\text{K}]_i]/[\text{K}]_o)^{n_{\text{K}}} \quad (1)$$

The effect of varying ion gradients on the glutamate transport current reversal potential was examined to determine the flux

coupling coefficient n for each ion. Increasing extracellular concentrations of L-glutamate resulted in a positive shift of the transport zero-current potential (Fig. 2*f*). Similar results were obtained for increasing H^+ and Na^+ concentrations, whereas increasing the K^+ concentration resulted in a negative shift of the reversal potential (Fig. 3). The reversal potential altered by 31.8 ± 1.5 , 25.2 ± 3.9 , 100.8 ± 3.4 , and $-31.3 \pm 0.9\ \text{mV}$ for each tenfold change in the concentration gradients of glutamate, H^+ , Na^+ and K^+ , respectively ($N = 5-7$; Fig. 3 insets). Notably, the reversal potential changes caused by altering the potassium and glutamate gradients were in the opposite direction to those expected for independent (uncoupled) flux of these ions. Defining a thermodynamic reaction cycle as translocation of one glutamate ion, the ratio of the slopes of the plots of the reversal-potential changes relative to glutamate gives the ratio of the coupling coefficients from equation (1) as $1\ \text{Glu}^- : 0.79\ \text{H}^+ : 3.17\ \text{Na}^+ : 0.98\ \text{K}^+$. The concentration dependencies of the respective unidirectional forward and reverse transport currents on $[\text{Glu}]_o$, $[\text{H}^+]_o$ and $[\text{K}^+]_o$ were well described by Hill equations with coefficient n not significantly different from 1, whereas the dependence of the forward transport current on $[\text{Na}^+]_o$ exhibited a Hill coefficient of 2.25 ± 0.12 ($N = 3$; Fig. 3).

The zero-flux equation (1) was used to predict the external glutamate concentration at equilibrium as a function of membrane potential, with ion concentration gradients similar to those found during ischaemia^{18,19} (Fig. 4*a*). Predictions of a model involving cotransport of one glutamate ion with three Na^+ ions and one H^+ ion, and countertransport of one K^+ ion were in good agreement with the measured reversal potentials (Fig. 4*a*, solid line), whereas the corresponding predictions for flux coupling with two rather than three Na^+ ions did not match the data (Fig. 4*a*, dashed line). A stoichiometry with three Na^+ ions was also found to be consistent with measurements of charge transfer during uptake of $100\ \mu\text{M}$ $[\text{H}^3]\text{L}$ -glutamate under voltage clamp

FIG. 3 External ion and glutamate concentration-dependence of forward (a, b and d; $V_m = -60$ mV) and reverse (c; $V_m = 20$ mV) transport. Data points (mean $\pm$ s.e.m.; $N = 3-5$) were fitted by least-square to $I/I_{\max} = [ion]^n / ([ion]^n + K^n)$ with $n = 1$ (panels a-c) or $n = 2.3$ (d). Affinity constants were 27 μ M (Glu), 26 nM (H^+), 46 mM (Na^+), and 17 mM (K^+). Kinetic parameters for forward transport (a, b and d) were measured in normal Ringer solution; in b and d, 10 μ M glutamate was present. K^+ -dependence of reverse transport (c) was measured in the absence of external glutamate with $[Na^+]_o = 60$ mM. Insets show transport current reversal potential shifts as a function of indicated external ion concentration. Points represent mean $\pm$ s.e.m. ($N = 5-9$ oocytes). From the slopes, the reversal potentials is seen to vary by 31.8, 25.2, -31.3 and 100.8 mV per tenfold change in concentration of the indicated ion, in a-d, respectively. Except where explicitly varied, the external concentrations were: $[Glu^-]_o = 10 \mu$ M; $[H^+]_o = 32$ nM; $[Na^+]_o = 60$ mM; $[K^+]_o = 40$ mM.

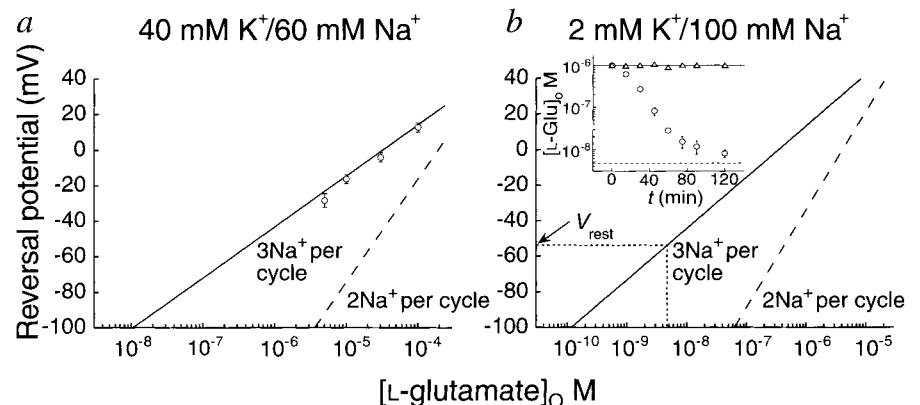


(2.04 ± 0.01 charges per molecule, $V_m = -30$ mV; $N = 7$). The coupling stoichiometry was further tested with ion gradients near physiological values by examining the ability of the transporter to concentrate L-glutamate from normal extracellular Ringer solution. After addition of 1 μ M L-glutamate, the concentration decayed to a roughly steady-state level of less than 10 nM within 2 hours, corresponding to a transmembrane concentration gradient exceeding 10^6 (Fig. 4b, inset). This glutamate concentration gradient is consistent with a flux coupling to three Na^+ ions per cycle but could not be generated with two Na^+ ions per cycle (Fig. 4b).

With a flux-coupling model of two Na^+ ions per glutamate molecule, the lower limit of extracellular glutamate in the brain

has been estimated to be $\sim 0.6 \mu$ M at -90 mV in physiological conditions¹. In the same conditions, flux coupling with three Na^+ per cycle would reduce this concentration to ~ 2 nM. Estimates of global ambient $[Glu]_o$ obtained using dialysis probes in brain tissue have yielded values in the submicromolar to low micromolar concentration range^{20,21}. Because these values are far from equilibrium, concentrations of glutamate in synapses and other regions with large numbers of transporters^{22,23} are expected to be lower. The existence of local glutamate concentration gradients in areas of restricted diffusion would be consistent with the observation of epileptic activity induced by antisense downregulation of the neuronal glutamate transporter in the absence of a change in mean extracellular glutamate levels²¹. Because ambient global

FIG. 4 a, Predictions of the zero-flux equation (1) for transport reversal as a function of extracellular $[L-Glu]$ in ischaemic conditions (60 mM Na^+ /40 mM K^+) compared with measured reversal potentials (data from Fig. 3a; mean $\pm$ s.e.m., $N = 5$). Solid line represents the flux predicted for $3Na^+ : 1H^+$ cotransport and counter-transport of $1K^+$ per transport cycle; dashed line represents the same equation with coupling of 2 rather than 3 Na^+ ions. b, Theoretical glutamate transport reversal potentials predicted for normal ion gradients (100 Na^+ ; 2 K^+) with solid ($n_{Na} = 3$) and dashed ($n_{Na} = 2$) lines as in a. Inset shows time course of external glutamate concentration decay after addition of 1 μ M L-glutamate to normal Ringer solution that was bathing oocytes expressing EAAT3 (circles)



or uninjected controls (triangles). Dotted line shows theoretical lower limit of 4.6 nM for mean resting potential (-54 ± 5 mV; $N = 4$).

glutamate is in a concentration range in which tonic activation and desensitization of pre- and postsynaptic glutamate receptors occurs^{24–26}, transporter localization as well as modulation of glutamate uptake rates by second messengers^{27–30} could influence receptor activity. The flux coupling of three rather than two sodium ions also has implications for glutamate transport when the potassium gradient is perturbed, for example, during transient ischaemia or intense neural activity, as the dominant contribution of the sodium gradient to the free-energy change will tend to maintain a thermodynamic driving force favouring forward transport of glutamate. □

Methods

Transporter expression and fluorescence measurements. Capped messenger RNA was injected into stage V–VI oocytes after transcribing linearized plasmid containing the coding region of EAAT3 (ref. 12) (gift from S. G. Amara and J. L. Arriza). Intracellular pH changes were recorded under voltage clamp conditions by monitoring 2',7'-bis-(2-carboxyethyl)-5-carboxyfluorescein (BCECF) fluorescence intensity using an inverted fluorescence microscope (Nikon Diaphot) with a $\times 20$ quartz objective and a P100S photomultiplier. Oocytes were loaded for 1 h with the acetomethoxyester derivative of BCECF (1 μ M BCECF-AM; Molecular Probes). The objective was focused on the vegetal pole and the excitation wavelength was alternated between 440 and 490 nm while monitoring emission intensity at 540 nm. Fluorescence signals were >100 times larger than background (unloaded oocytes) and signals at 440 nm excitation did not change during the duration of the recordings shown. In matched comparisons in 7 cells, the fluorescence change during cysteine transport was $5.4 \pm 1.6\%$ of that observed with glutamate.

Electrophysiological recordings and modelling. Transport currents were measured with a two-microelectrode voltage-clamp circuit. Except for Fig. 1, recordings were made with Cl[−]-free Ringer (gluconate substitution) with internal Cl[−] dialysed as described⁷. Solutions normally contained 100 mM Na⁺, 1 mM Mg²⁺, 1.8 mM Ca²⁺, 2 mM K⁺, 4 mM Ba²⁺ and 5 mM HEPES. When monovalent cation concentrations were varied, the total cation concentration was held constant with Tris substitution. Reversal potential shifts of the transport current were determined from the current blocked by kainate (RBI; Tris salt) with varying external concentrations of the tested ionic species. Modelling of the transport–voltage relationship was done using equation (1), assuming intracellular values of [Na⁺]_i = 10 mM, [K⁺]_i = 70 mM, pH_i = 7.3, and excitatory amino-acid concentration of 12 mM (ref. 14).

Glutamate flux measurements. Uptake of glutamate was measured by assaying aliquots of solution from multiwell plates containing oocytes expressing EAAT3 or uninjected controls (15 oocytes per ml Ringer) at indicated times after addition of 1 μ M L-glutamate. Glutamate concentrations were determined by measuring current responses of oocytes expressing a high-affinity mutant Y373W EAAT3 transporter ($K_m = 0.5 \pm 0.08 \mu$ M; B. I. Kanner, N.Z. and M.P.K., manuscript in preparation) to pressure ejection of 100 μ l aliquots and comparing responses to known concentration standards. Measurement of charge transfer per molecule of glutamate was determined as described⁷ by superfusing [³H]-L-glutamate (Amersham) onto Cl[−]-depleted oocytes voltage clamped at -30 mV in Cl[−]-free Ringer. Measurements in the presence of normal Cl[−] gradients with 4 different batches of [³H]-L-glutamate in 9 groups of oocytes voltage clamped at E_{Cl} (~ -25 mV) gave similar results (2.16 ± 0.03 charges per molecule Glu; $N = 30$).

Received 20 June; accepted 6 September 1996.

- Bouvier, M., Szatkowski, M., Amato, A. & Attwell, D. *Nature* **360**, 471–474 (1992).
- Kanner, B. I. & Sharon, I. *Biochemistry* **17**, 3949–3953 (1978).
- Barbour, B., Brew, H. & Attwell, D. *Nature* **335**, 433–435 (1988).
- Erecinska, M., Wantorsky, D. & Wilson, D. F. *J. Biol. Chem.* **258**, 9069–9077 (1983).
- Nelson, P. J., Dean, G. E., Aronson, P. S. & Rudnick, G. *Biochemistry* **22**, 5459–5463 (1983).
- Fairman, W. A., Vandenberg, R. J., Arriza, J. L., Kavanaugh, M. P. & Amara, S. G. *Nature* **375**, 599–603 (1995).
- Wadiche, J. I., Amara, S. G. & Kavanaugh, M. P. *Neuron* **15**, 721–728 (1995).
- Larsson, H. P., Picaud, S. A., Werblin, F. S. & Lecar, H. *Biophys. J.* **70**, 733–742 (1996).
- Eliasof, S. & Jahr, C. E. *Proc. Natl. Acad. Sci. USA* **93**, 4153–4158 (1996).
- Billups, B., Rossi, D. & Attwell, D. *J. Neurosci.* (in the press).
- Kanai, Y. & Hediger, M. A. *Nature* **360**, 467–471 (1992).
- Arriza, J. L. et al. *J. Neurosci.* **14**, 5559–5569 (1994).
- Billups, B. & Attwell, D. *Nature* **379**, 171–174 (1996).
- Zerangue, N. & Kavanaugh, M. P. *J. Neurosci.* **17**, 419–423 (1996).
- Webb, D. J. & Nuccitelli, R. *J. Cell Biol.* **91**, 562–567 (1981).
- Burger, P. M. et al. *Neuron* **3**, 715–720 (1989).
- Szatkowski, M., Barbour, B. & Attwell, D. *Nature* **348**, 443–446 (1990).
- Hansen, A. J. *Physiol. Rev.* **65**, 101–148 (1985).
- Szatkowski, M. & Attwell, D. *Trends Neurosci.* **17**, 359–365 (1994).
- Benveniste, H., Drejer, J., Schousboe, A. & Diemer, N. H. *J. Neurochem.* **43**, 1369–1374 (1984).
- Rothstein, J. D. et al. *Neuron* **16**, 675–686 (1996).
- Chaudry, F. A. et al. *Neuron* **15**, 711–720 (1995).
- Rothstein, J. D. et al. *Neuron* **13**, 713–725 (1994).

- Saugstad, J. A., Segerson, T. P. & Westbrook, G. L. in *Excitatory Amino Acids and Synaptic Transmission* (eds Wheal, H. & Thomson, A.) 77–85 (Academic, London, 1995).
- Trussel, L. O. & Fischbach, G. D. *Neuron* **3**, 209–218 (1989).
- Patneau, D. K. & Mayer, M. L. *J. Neurosci.* **10**, 2385–2399 (1990).
- Chan, P. H., Kerlan, R. & Fishman, R. A. *J. Neurochem.* **40**, 309–316 (1983).
- Barbour, B., Szatkowski, M., Ingledew, N. & Attwell, D. *Nature* **342**, 918–920 (1989).
- Zerangue, N., Arriza, J. L., Amara, S. G. & Kavanaugh, M. P. *J. Biol. Chem.* **270**, 6433–6435 (1995).
- Volterra, A., Trotti, D., Tromba, C., Floridi, S. & Racagni, G. *J. Neurosci.* **14**, 2924–2932 (1994).

ACKNOWLEDGEMENTS. We thank D. Attwell for comments on the manuscript. This work was supported by the NIH.

CORRESPONDENCE and requests for materials should be addressed to M.P.K. (e-mail: kavanaugh@ohsu.edu).

Cyclophilin-related protein RanBP2 acts as chaperone for red/green opsin

Paulo A. Ferreira, Tomoko A. Nakayama*, William L. Pak† & Gabriel H. Travis

Department of Psychiatry and Program in Neuroscience, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75235-9111, USA

* Department of Biostructure and Function, The University of Connecticut Health Center, 263 Farmington Avenue, Farmington, Connecticut 06030-3705, USA

† Department of Biological Sciences, Lilly Hall of Life Sciences, Purdue University, West Lafayette, Indiana 47907, USA

CYCLOPHILINS are ubiquitous and abundant proteins that exhibit peptidyl prolyl *cis*–*trans* isomerization (PPIase) activity *in vitro*^{1,2}. Their functions *in vivo*, however, are not well understood. Two new retinal cyclophilin isoforms, types I and II, are highly expressed in cone photoreceptors of the vertebrate retina³. Type-II cyclophilin is identical to RanBP2, a large protein that binds the GTPase Ran^{4,5}. Here we report that two contiguous domains in RanBP2, Ran-binding domain 4 (RBD4) and cyclophilin, act in concert as a chaperone for the opsin molecule of the red/green-sensitive visual pigment of a dichromatic vertebrate. In *Drosophila*, the cyclophilin NinaA^{6,7} is expressed in all photoreceptors⁸ and is required for the expression of only a subset of opsins^{8,9}. The molecular basis of these photoreceptor class-specific effects and the functions of NinaA and other cyclophilins *in vivo* remain unclear¹⁰. Unlike NinaA, which forms a stable complex with opsin from retinal cells R1–6¹¹, we find that the cyclophilin domain of RanBP2 does not bind opsin directly; rather, it augments and stabilizes the interaction between red/green (R/G) opsin and the RBD4 domain. This involves a cyclophilin-mediated modification of R/G opsin, possibly involving proline isomerization. The RBD4–cyclophilin supradomain of RanBP2, therefore, is a form of vertebrate chaperone of defined substrate specificity, which may be involved in the processing and/or transport of long-wavelength opsin in cone photoreceptor cells.

To investigate the function of the new vertebrate retinal cyclophilins in photoreceptor cells, we expressed a series of complementary DNA constructs encoding different fragments of RBD4 and cyclophilin fused to a glutathione-S-transferase (GST) carrier protein (Fig. 1a). These fusion proteins were incubated with detergent-solubilized bovine retinal extracts, and their coprecipitating substrates were analysed by immunoblotting. We used a candidate-protein approach to see which photoreceptor proteins may bind to RBD4–cyclophilin and subfragments thereof (Fig. 1b–f). Immunoblot analysis of retinal co-precipitates from incubation reactions with an antiserum against R/G opsin¹² showed high-affinity GTP-independent binding of the opsin to the complete RBD4–cyclophilin construct

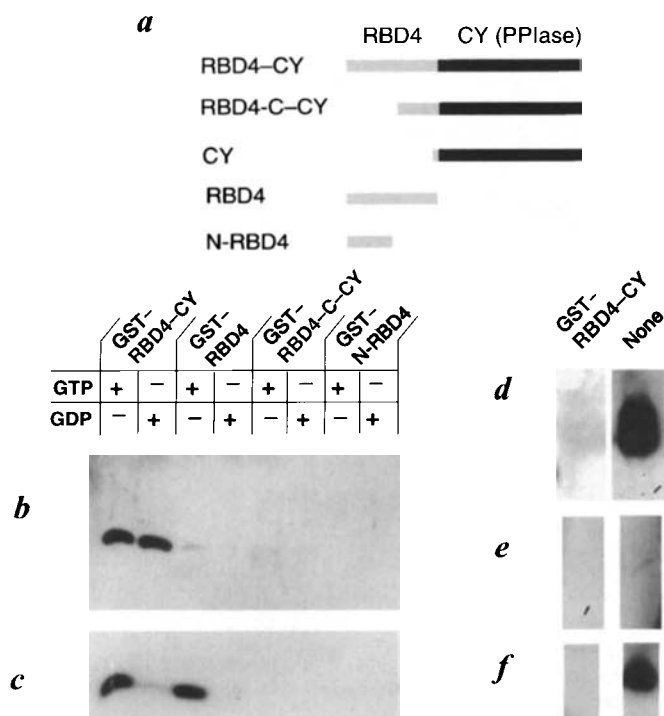


FIG. 1 The RBD4-cyclophilin (CY) supradomain binds selectively to R/G opsin. **a**, Schematic drawings of the RBD4-cyclophilin constructs used in binding assays. RBD4-CY contains the entire Ran-binding domain 4 (RBD4) and the cyclophilin domain (CY), with PPIase catalytic site³. RBD4 C-CY contains the carboxy-terminal half of RBD4 plus cyclophilin. CY contains the cyclophilin domain only. RBD4 contains the entire RBD4. N-RBD4 contains the amino-terminal half of RBD4. **b**, **c**, Western blot analysis of glutathione-S-agarose coprecipitates from incubation reactions of retinal extracts with GST-fused RBD4-cyclophilin and several subfragments using a polyclonal antibody against R/G opsin (**b**) and Ran GTPase (**c**). **d**, **e**, **f**, Western blot analysis of retinal coprecipitates with GST-fused RBD4-cyclophilin (first lane) and an aliquot of total retinal extract (second lane) with antiserum against rhodopsin (**d**), blue cone opsin (**e**) and rds/peripherin (**f**). The faint signal in the aliquot of total retinal extract (second lane) in (**e**) reflects the rareness of blue cone photoreceptors in bovine retina.

(Fig. 1b). Binding to RBD4 alone was greatly reduced, and we saw no binding to either cyclophilin, subfragments of RBD4, or the GST fusion partner alone (Figs 1b and 2a). In contrast, the binding of Ran-GTPase to RBD4 was GTP-dependent and unaffected by the absence of cyclophilin (Fig. 1c). No binding was observed with the closely related visual pigments, blue opsin¹³ and rhodopsin¹⁴ (Fig. 1d, e), nor with the outer segment disc protein, rds/peripherin¹⁵ (Fig. 1f). Thus, the interaction between RBD4 and R/G opsin is highly specific and significantly augmented by cyclophilin.

Immunoblot analysis with antiserum against R/G opsin¹² on an aliquot of the incubation reaction containing either no added protein or the GST fusion partner alone, revealed a single immunoreactive band with an apparent molecular mass of 49.5K (Fig. 2b). This is close to the predicted relative molecular mass of human red opsin (47.1K)¹³. However, in the presence of RBD4-cyclophilin, either fused to GST or unfused, we detected a second immunoreactive band of 51K. In binding studies, only this reduced electrophoretic mobility form of R/G opsin bound to RBD4-cyclophilin (Fig. 2a), implying that it modifies R/G opsin in some way, and that only this modified form is competent to bind RBD4-cyclophilin. The nature of this modification was not clear. Treatment with the broad-spectrum phosphatase calcineurin (Fig. 2b), or N-glycosidase F (data not shown) had no effect on the mobility of R/G opsin, indicating that RBD4 does not selectively bind a phosphorylated or N-glycosylated form. We

saw no difference in the two forms of R/G opsin at the start or after three hours of incubation with RBD4-cyclophilin, ruling out protection against proteolysis as an explanation for the mobility shift. As before, only the reduced-mobility form of R/G opsin was associated with RBD4-cyclophilin (Fig. 2c).

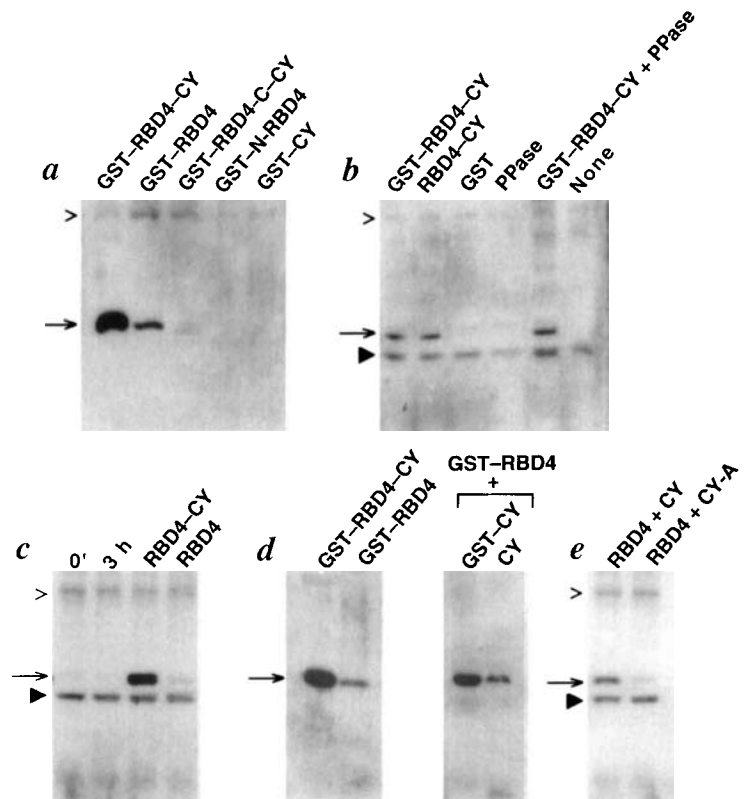
The reduced electrophoretic mobility form of R/G opsin might be caused by isomerization of one or several proline residues, catalysed by the PPIase activity of cyclophilin. According to this model, prolyl isomerization and binding of R/G opsin to RBD4 may represent sequential transformations induced by the adjacent cyclophilin and RBD4 domains. One prediction of this model is that modification and binding of R/G opsin may be induced *trans* by the RBD4 and cyclophilin domains as separate polypeptide chains. To test this, we did binding assays on bovine retinal extracts with separate RBD4 and cyclophilin protein fragment, independently fused to GST (Fig. 2d, left panel). There was increased binding of R/G opsin to RBD4 in the presence of cyclophilin (Fig. 2d, left panel, first lane). In contrast, when we incubated GST-fused RBD4 with unfused cyclophilin polypeptide, which was later removed during the brief precipitation washes, there was no augmentation of binding (Fig. 2d, left panel, second lane). Removal of cyclophilin immediately after production of the putative PPIase-induced, reduced electrophoretic mobility form of R/G opsin (Fig. 2e, first lane), lead to a subsequent decrease in the amount of this to levels as when cyclophilin had not been added (Fig. 2d). To investigate the specificity of cyclophilin for R/G opsin, we replaced the polypeptide in the retinal extract with the related protein, cyclophilin A (65% identical³), which has higher PPIase activity on a standard peptidyl-prolyl substrate³. Here, there was no conversion of R/G opsin to its reduced-mobility, binding-competent form, in contrast to significant conversion with the free cyclophilin polypeptide (Fig. 2e), indicating that cyclophilin is functioning as a specific chaperone and foldase for R/G opsin in its interaction with RBD4.

It is also possible that the changes in R/G opsin are due to some other stable PPIase-dependent modification, but this is inconsistent with our observation that removal of unfused cyclophilin from the incubation reaction resulted in a rapid reduction of R/G opsin-binding to RBD4 (Fig. 2d). On the contrary, if the cyclophilin-PPIase domain were functioning (either in *cis* or in *trans*) to render R/G opsin susceptible to some other stable post-translational modification, we would have expected augmented modification of R/G opsin and increased binding to RBD4.

To test whether binding of R/G opsin to RBD4 and its associated change in electrophoretic mobility are dependent upon cyclophilin PPIase activity, we analysed the PPIase catalytic domain of cyclophilin by mutagenesis. Several residues within the catalytic domain of the retinal cyclophilin are highly conserved among all cyclophilins³. Enzyme kinetic and crystallographic analysis have shown that arginine at residue 55 and phenylalanine at residue 60 are critical for the PPIase activity of cyclophilin A¹⁶⁻¹⁸. We introduced the equivalent mutation, Arg substituted by Ala at position 143(R143A), and the double mutation, R143A with F148A, into the cognate PPIase domain of RBD4-cyclophilin. These substitutions dramatically decreased binding of R/G opsin to RBD4 (Fig. 3a), and conversion of R/G opsin to its reduced electrophoretic-mobility form (Fig. 3b). Thus, the observed changes in R/G opsin seem to be dependent upon the PPIase activity of cyclophilin.

In addition, conversion of R/G opsin to its RBD4-cyclophilin binding-competent form was associated with another change in its physical properties. The 51K isoform of R/G opsin, affinity-purified by binding to RBD4-cyclophilin, showed no propensity to self-aggregate with SDS-PAGE (polyacrylamide gel electrophoresis). In contrast, when the entire incubation mixture was heated, followed by SDS-PAGE, the typical self-aggregation ladder pattern of opsin was apparent^{11,19,20} (Fig. 3b), as the majority of R/G opsin in this mixture is not in its RBD4-cyclophilin binding-competent form. Similar results were observed when we co-expressed GST-RBD4-cyclophilin with

FIG. 2 Binding of the RBD4–cyclophilin supradomain and component subfragments to a modified form of R/G opsin. The cyclophilin domain of RBD4–CY is a specific chaperone for R/G opsin. Western blot analysis of glutathione-S-agarose coprecipitates from incubation reaction extracts (a) and aliquots of incubation reactions (b) with indicated GST-fusion proteins, using a polyclonal antibody against R/G opsin. (c) Western blot analysis of aliquots of incubation reaction extracts at time 0 (lane 1) and after 3 h with no added protein (lane 2), with added RBD4–cyclophilin (lane 3) and with added RBD4 alone (lane 4). d, Western blot analysis of glutathione-S-agarose coprecipitates from incubation reactions with GST-fused RBD4–cyclophilin and RBD4 in the presence of added GST-fused or unfused cyclophilin, using antiserum against R/G opsin. No augmentation of binding to RBD4 is seen with the addition of unfused cyclophilin, which is removed during precipitation washes (right panel, second lane). e, Western blot analysis of aliquots of incubation reactions containing RBD4 plus cyclophilin or cyclophilin A (CY-A), using a polyclonal antibody against R/G opsin. Filled arrowhead, arrow and open arrowhead denote native/nonbinding form, binding form and dimers of R/G opsin, respectively.



red opsin in *Escherichia coli* (P.A.F., T.A.N. and G.H.T., manuscript submitted).

To investigate the functional role of RBD4–cyclophilin on red opsin *in vivo*, we co-expressed the red opsin apoprotein with either RBD4–cyclophilin, RBD4 and cyclophilin in COS-1 cells. Only about 20% of human red opsin expressed in COS-1 cells bind 11-*cis*-retinal to generate functional pigment. This is very low in comparison to the nearly complete generation of functional rhodopsin from rod opsin expressed in COS-1 cells³⁰. Co-expression of RBD4–cyclophilin with red opsin resulted in a doubling in functional red pigment during a one-hour incubation with 11-*cis*-retinal (Fig. 4). In contrast, co-expression with either RBD4 or cyclophilin led to small or no increase, respectively, of functional pigment formation (Fig. 4). Thus, expression of RBD4 with cyclophilin enhances formation of functional pigment in COS-1 cells. In *Drosophila* photoreceptors, at least ten different gene products, NinaA–J, are required for normal levels of functional opsin pigment^{3,21,22}. The *Drosophila* NinaA protein has been shown to be required for proper targeting of opsins in R1–6, to the microvillar membrane of the rhabdomeres^{8,9}. In contrast to RanBP2, NinaA does not contain a Ran-binding domain. As cyclophilin can act in *trans* to RBD4, it is possible that in *Drosophila*, the cyclophilin and Ran-binding functions are subserved by separate proteins. These *in vivo* observations agree with our observations *in vitro*, and further indicate that the chaperone and foldase activity of RBD4–cyclophilin is necessary for production of functional pigment.

Several conclusions can be drawn from our observations. The selective binding of R/G opsin by RBD4 requires cyclophilin. This binding seems to be dependent on the PPIase-activity of cyclophilin, and thus may be associated with isomerization of one or more proline residues within R/G opsin, reflected by changes in the electrophoretic mobility and low propensity of the RBD4-binding competent isoform to form self-aggregates in the presence of SDS. Although the cyclophilin domain can assist RBD4-binding of R/G opsin both in *cis* and in *trans*, these two domains must be in close physical proximity (bound to the same glutathione–agarose bead, Fig. 2d), implying that cyclophilin chaperones, the interac-

tion between RBD4 and R/G opsin. The interactions between cyclophilin and R/G opsin, and/or between cyclophilin and RBD4, are specific for this form of cyclophilin. Furthermore, the RBD4–cyclophilin supradomain of RanBP2 is required *in vivo* to enhance the production of functional red pigment probably by the specific

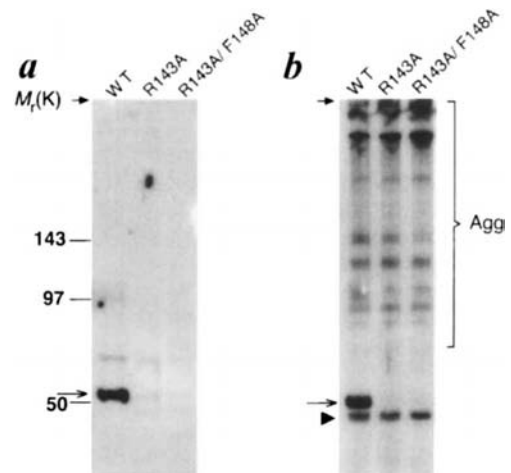


FIG. 3 Mutations in the PPIase domain of RBD4–cyclophilin inhibit the modification and binding of R/G opsin. Western blot analysis of glutathione-S-agarose coprecipitates from incubation reactions (a) and aliquots of incubation reactions (b) with wild-type (WT) and mutated GST-fused RBD4–cyclophilin, using a polyclonal antibody against R/G opsin. The single-point mutation, R143A, and the double mutation, R143A/F148A, in the PPIase catalytic site of RBD4–cyclophilin significantly decreases the binding of R/G opsin to RBD4–cyclophilin (a) and abolishes the conversion of R/G opsin to the reduced-mobility, binding-competent form (b). Amino acid numbering corresponds to type I cyclophilin³. Filled arrow, open arrow, and arrowhead represent origin of the gel, binding-competent, and native non-binding forms of R/G opsin, respectively. Agg, higher molecular mass aggregates of R/G opsin. Position of protein size standards is indicated to the right.

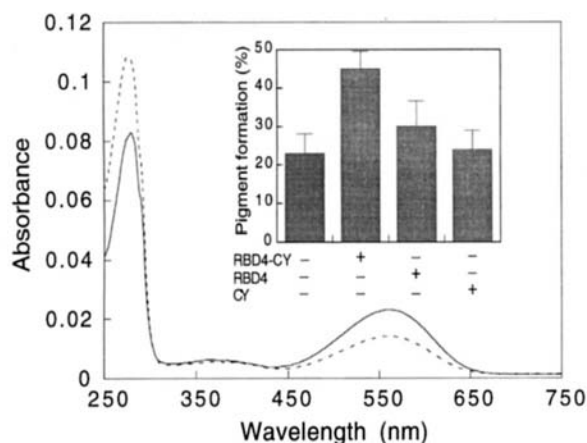


FIG. 4 Absorption spectra of human red pigment expressed in COS cells with (solid line) and without (dotted line) RBD4-cyclophilin. Co-expression of red opsin with RBD4-cyclophilin roughly doubled pigment formation over singly expressed red opsin. In contrast, co-expression of red opsin with RBD4 and cyclophilin led to a slight and no increase in the extent of pigment formation, respectively (inset).

chaperone and foldase activity of RBD4-cyclophilin. Given the high concentration of opsin visual pigments in photoreceptors (90% of outer segment disk protein)^{19,23}, the processing of these membrane proteins may require a special machinery.

The function of Ran binding to RanBP2 remains unclear. One possibility is that RanBP2 may be a tissue-restricted effector of Ran, which may couple the Ran-dependent subcellular shuttling of certain messenger RNAs (for example, R/G opsin) from the nucleus to the cytoplasm^{24,25} with the RanBP2-dependent processing of the R/G opsin. Ran function is probably dependent upon the nuclear guanine-nucleotide-exchange factor RCC1, which catalyses the conversion of Ran-GDP to Ran-GTP²⁶. A ubiquitous and putatively cytosolic RCC1 homologue, RPGR (retinitis pigmentosa GTPase regulator), has been identified and cloned²⁷⁻²⁹. It is possible that mutations in the RPGR gene could impair the coupling of Ran with its effector targets, some of which, like RanBP2, may exhibit tissue-restricted localization. Mutations in the RPGR gene have been identified that lead to the very severe retina-specific phenotype X-linked retinitis pigmentosa (RP3)²⁷⁻²⁹. Thus, RanBP2-dependent processing and targeting of opsin may be dependent on Ran and RPGR. For this reason, the gene for RanBP2 is a strong candidate for involvement in additional inherited blindness diseases. □

Methods

Preparation of GST fusion constructs. The RBD4 protein fusion construct was prepared by deleting the NcoI restriction fragment of RBD4-cyclophilin, followed by re-ligation. The preparation of all other constructs and purification of their encoded proteins has been described³. GST-fused proteins were purified through competitive elution with 10 mM reduced glutathione. Proteins were concentrated and washed 3–4 times in 50 mM Tris-HCl (pH 7.5), 0.1% β -mercaptoethanol using Centricon concentrators (Amicon). Protein quantification was done by the Bio-Rad protein assay method.

Preparation of retinal extracts and binding assays. Fresh bovine retinas were homogenized (1:3 w/v) at 4 °C in 0.75% CHAPS, 20 mM Tris-HCl (pH 6.8), 250 mM NaCl, 2 mM β -mercaptoethanol, 0.02% Na₂S₂O₈, 5% glycerol and protease inhibitors. After centrifugation at 8,000g at 4 °C for 30 min, supernatants were preincubated with glutathione-S-agarose (1:20 dilution of the agarose) and GST (1–5% w/v) at 4 °C for 1 h. Incubation mixtures consisted of 80 μ l of retinal extract (~4 mg of total protein), 2.2 μ M of GST fusion protein, 0.25 μ M of GTP γ S or GDP β S (when applicable). After incubation on a shaker for 1.5 h at 4 °C, 60 μ l of 50% (w/v) glutathione-S-agarose beads in incubation buffer (50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 2 mM CaCl₂, 2 mM MgCl₂ and 0.5% CHAPS) was added to incubation reactions for 1 h. For the binding assays, the beads were briefly microcentrifuged (3 s), washed (3 times) with 0.5–1 ml of incubation buffer containing 0.2% Triton X-100 and resuspended in SDS sample buffer. Boiled co-precipitates were loaded for 10% SDS-PAGE (13.5% in Fig. 1b,

c) and electroblotted onto PVDF membranes (Millipore). In other cases, aliquots of incubation reactions mixtures (~120 μ g of total protein) were loaded for SDS-PAGE. Development of the blots with rabbit anti-R/G opsin (JH 492, 1:2,500); anti-Ran/TC4 (1:1,500); monoclonal anti-rhodopsin (4D2, 1:60); anti-blue opsin (JH 455, 1:5,000); or anti-D2P4 peptide from human rds/peripherin (1:1,000) was done using the chemiluminescent kit (Kirkgaard & Perry Laboratories, Inc.) as described by the manufacturer. The blocking solution contained 10% glycerol, 2% non-fat dry milk (BioRad), 1 $\times$ PBS (pH 7.4) and both primary and secondary antibodies were diluted in 1:5 blocking solution. Phosphatase treatment was done as described by the manufacturer instructions (Boehringer Mannheim). Equimolar concentrations of cyclophilin A (Boehringer Mannheim) and retinal cyclophilin were used in Fig. 2e.

Construction of RBD4-cyclophilin PPLase-deficient mutants. For construction of PPLase active-site mutants, we made two pairs of purified complementary and mutagenized synthetic oligonucleotides: 5'-CCATATCCAC GCAGTGATTCAGATTTTGTTCGCAAGGGGAGAT and 3'-TCGAGGTATAAGGTGTC GTCATAAGGTCTAAAACAAACGGTTCCTCTCTA; and 5'-CCATATCCACGCAGTG ATTCCAGATGCAGTTTTCGCAAGGGGAGAT and 3'-TCGAGGTATAAGGTGCGCTCACT AAGGTCTACGTCAACACGGTTCCTCTCTA encoding the point mutations of R143A and R143A/F148A, respectively. Oligonucleotide pairs were mixed in equimolar amounts, heated to 65 °C for 5 min, and slowly cooled to 30 °C over 30 min. The annealed oligonucleotides were ligated into SacI-EcoRV 350-bp cyclophilin cDNA fragment, phosphorylated, and ligated into the SacI-digested, dephosphorylated GST-RBD4-cyclophilin construct.

Co-expression of red opsin and RBD4-cyclophilin constructs in COS cells. To facilitate purification of the rhodopsin carboxy-terminal 14 amino acids containing the monoclonal antibody 1D4 epitope were added to the red pigment C terminus by cloning the PCR modified red opsin cDNA into the 5.1-kb EcoRI-SalI fragment of the rhodopsin expression vector pMT3. A separate expression vector for RBD4-cyclophilin, RBD4, or cyclophilin was constructed by cloning the corresponding cDNA into the 5.0-kb EcoRI-NotI fragment of pMT3. Red opsin was transiently expressed singly or in combination with RBD4-cyclophilin, RBD4, or cyclophilin expressed at similar levels. Transfection, pigment generation, and purification of the red pigment were previously described³⁰ except the buffer used for the purification was 40 mM HEPES, pH 6.5, 130 mM NaCl, and 20% glycerol. Spectroscopic measurements were done using a Perkin-Elmer Lambda 14 UV/Vis spectrophotometer on purified red pigment. The extent of pigment formation was estimated by using $\epsilon_{560\text{ nm}}$ and $\epsilon_{280\text{ nm}}$ values of 50,000 cm² M⁻¹ and 95,000 cm² M⁻¹, respectively.

Received 15 July; accepted 21 August 1996.

- Fisher, G., Wittmann-Liebold, B., Lang, K., Kieffhaber, T. & Schmid, F. *Nature* **337**, 476–478 (1989).
- Bergsma, D. et al. *J. Biol. Chem.* **266**, 23204–23214 (1991).
- Ferreira, P., Hom, J. & Pak, W. *J. Biol. Chem.* **270**, 23179–23188 (1995).
- Yokoyama, N. et al. *Nature* **376**, 184–188 (1995).
- Wu, J., Manutis, M., Kraemer, D., Blobel, G. & Coutavas, E. *J. Biol. Chem.* **270**, 14209–14213 (1995).
- Schneuwly, S. et al. *Proc. Natl Acad. Sci. USA* **86**, 5390–5394 (1989).
- Shieh, B.-H., Stammers, M., Seavell, S., Harris, G. & Zuker, C. *Nature* **338**, 67–70 (1989).
- Stammes, A., Shieh, B.-H., Chuman, L., Harris, L. & Zuker, C. *Cell* **65**, 219–227 (1991).
- Larivée, D., Conrad, S., Stephenson, R. & Pak, W. *J. Gen. Physiol.* **78**, 521–545 (1981).
- Ferreira, P. & Pak, W. In *Degenerative Diseases of the Retina* (eds Angerson, J., La Vail, J. & Hollyfield, J.) 263–274 (Plenum, New York, 1995).
- Baker, E., Colley, N. & Zuker, C. *EMBO J.* **13**, 4886–4895 (1994).
- Wang, Y. et al. *Neuron* **9**, 429–440 (1992).
- Nathans, J., Thomas, D. & Hogness, D. *Science* **232**, 193–202 (1986).
- Nathans, J. & Hogness, D. *Cell* **34**, 807–814 (1983).
- Travis, G., Sutcliffe, J. & Bok, D. *Neuron* **6**, 61–70 (1991).
- Zydowsky, L. et al. *Protein Sci.* **1**, 1092–1099 (1992).
- Kallen, H. & Walkinshaw, M. *FEBS Lett.* **300**, 286–290 (1992).
- Ke, H., Mayrose, D. & Cao, W. *Proc. Natl Acad. Sci. USA* **90**, 3324–3328 (1993).
- Kamps, K., De Grip, W. & Daemen, F. *Biochim. Biophys. Acta* **687**, 296–302 (1982).
- Walter, A., Shuster, T. & Farber, D. *Invest. Ophthalmol. Vis. Sci.* **27**, 1609–1614 (1986).
- Stephenson, R. S., O'Tousa, J., Scavarda, N. J., Randall, L. L. & Pak, W. L. In *The Biology of Photoreception* (eds Cosens, D. J. & Vince-Price, D.) 447–501 (Cambridge Univ. Press, 1983).
- Pak, W. L. *Invest. Ophthalmol. Vis. Sci.* **36**, 2340–2357 (1995).
- Molday, R. S. & Molday, L. L. *J. Cell Biol.* **105**, 2598–2601 (1987).
- Schlenstedt, G., Saavedra, C., Loeb, J., Cole, C. & Silver, P. *Proc. Natl Acad. Sci. USA* **92**, 225–229 (1995).
- Ren, M. et al. *Mol. Cell. Biol.* **15**, 2117–2124 (1995).
- Klebe, C., Bischoff, F. R., Ponstingl, H. & Wittinghofer, A. *Biochemistry* **34**, 639–647 (1995).
- Roepman, R. et al. *Hum. Mol. Genet.* **5**, 827–833 (1996).
- Roepman, R. et al. *Hum. Mol. Genet.* **5**, 1035–1041 (1996).
- Meindl, A. et al. *Nature Genet.* **13**, 35–42 (1992).
- Nakayama, T. A. & Khorana, H. G. *J. Biol. Chem.* **266**, 4269–4275 (1991).

ACKNOWLEDGEMENTS. We thank J. P. Albanesi for critical reading of the manuscript, M. Kung for technical assistance, A. Wright for a preprint of his paper on RPGR, J. Nathans for the h57 cDNA clone, JH 492 and JH 455 antisera, and helpful discussions. We thank R. Molday and P. D'Eustachio for the monoclonal 4D2 and anti-Ran/TC4 antibodies, respectively, and A. Mahrenholz for protein microsequencing. This work was funded by the NEI and Fight for Sight, Inc., Research Division of Prevent Blindness America, and an ACS institutional grant.

CORRESPONDENCE and requests for materials should be addressed to P.A.F. (e-mail: ferreira@utsw.swmed.edu).

Retrotransposon reverse-transcriptase-mediated repair of chromosomal breaks

Shu-Chun Teng, Bohye Kim & Abram Gabriel

Department of Molecular Biology and Biochemistry, Rutgers University, Piscataway, New Jersey 08855, USA

THE abundance of short and long interspersed nuclear sequences (SINEs and LINEs) and pseudogenes in eukaryotic genomes indicates that reverse transcriptase (RT)-mediated phenomena are important in genome evolution. However, the mechanisms involved in their spread are largely unknown. We have developed a selection system in the yeast *Saccharomyces cerevisiae* to test whether RT-mediated events could be linked to the repair of double-strand breaks (DSBs). Here we show that DSBs can be fixed by the insertion of complementary DNAs at the break site. In the presence of functional RT (from human L1, yeast Ty1 or *Crithidia* CRE1), and in the absence of homologous recombination, an HO endonuclease-induced DSB at the mating type (*MAT*) locus is the primary site at which a marked cDNA is observed among surviving cells. The structure and junctional sequences of these insertions suggest that repair occurs primarily by non-homologous recombination. Our data support a role for endogenous retroelements in the repair of chromosomal breaks.

RTs from a variety of sources are capable of synthesizing cDNAs in yeast¹⁻⁵. RT activity is detectable by an assay that depends on the expression of the *mhil3AI* marker gene, which contains an artificial intron inserted in the antisense orientation relative to *HIS3*. Histidine prototrophy requires transcription from a heterologous promoter in the antisense orientation, followed by splicing, reverse transcription, and insertion of the resulting cDNA into the genome^{6,7}. We initially examined the RT activity resulting from overexpression, in yeast, of the second open reading frame (ORF2) from a functional human LINE element (L1.2). We replaced the yeast retrotransposon Ty1 RT domain with the complete L1 ORF2, in a Ty1 construct driven by the yeast *GAL1* promoter⁸. The *mhil3AI* marker was inserted into the L1.2 3' untranslated region (UTR) (pGTyL1*mhil3AI*) (Fig. 1a). Galactose induction of cells containing pGTyL1*mhil3AI* results in overexpression of Ty1 virus-like particles (VLPs) containing both L1 RT and Ty1-L1-*his3* mRNA, and yields L1 RT-dependent histidine prototrophs (Fig. 1b). Deletion of 454 bases from the 3' end of the L1 segment (pGTyL1Δ454*mhil3AI*) results in ~100-fold increased frequency of histidine prototrophs (Fig. 1b) owing to increased antisense *HIS3* RNA stability (B.K. and A.G., unpublished data). Most cDNA insertions result from homologous recombination with pre-existing genomic Ty1 elements^{3,7}; the frequency of induced histidine prototrophy decreases ~100-fold in a *rad52* isogenic strain (data not shown).

To determine whether L1 RT-mediated phenomena could be linked to repair of chromosomal breaks, as previously proposed for LINE elements⁹, we took advantage of the yeast mating type (*MAT*) locus. In yeast, HO endonuclease makes a unique double-stranded cut near the Y-Z junction of the *MAT* locus. This break stimulates gene conversion of the site, using homologous donor sequences from the *HML* or *HMR* loci¹⁰. When homologous recombination is blocked (for example, in *rad52* strains), HO-induced DSBs are lethal, unless repaired by a non-homology-dependent mechanism (reviewed in ref. 11). We selected surviving histidine prototrophs following overexpression of both a *GAL*-HO plasmid and a derivative of the pGTyL1*mhil3AI* plasmid in a *rad52* strain (Fig. 2). Because insertion of *HIS3* cDNA at the break site would result in a genetic alteration of the *MAT* locus that heals the cut ends, we postulated that such events might be enriched in our selected population.

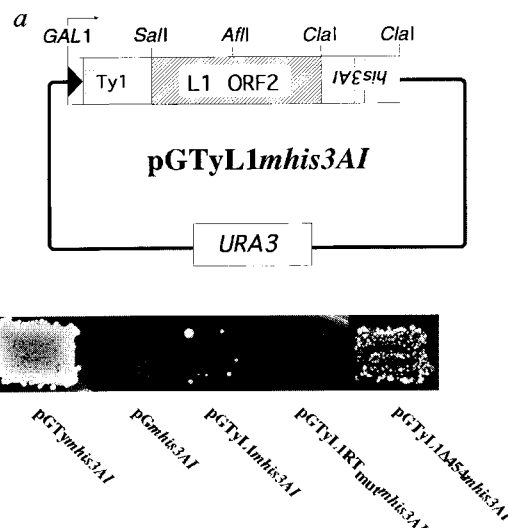


FIG. 1 L1 RT mediates histidine prototroph formation in yeast. a, Schematic representation of the prototypical plasmid used, containing a truncated galactose-inducible Ty1 element fused to L1 ORF2, and a downstream *his3AI* marker gene for cDNA detection. b, Patch assay demonstrating the ability of various constructs to create histidine prototrophs. pGTyL1*mhil3AI* contains a galactose-driven wild-type Ty1; histidine prototrophy results from transposition of the *HIS3*-marked Ty1 element. pG*mhil3AI* contains only the *mhil3AI* gene under the *GAL1* promoter⁷. pGTyL1*mhil3AI* contains the complete L1.2 ORF2, as in a. pGTyL1*RTmut**mhil3AI* is identical to pGTyL1*mhil3AI* except for a two-base substitution causing the substitution of a highly conserved aspartic acid residue in the L1 RT region⁸. pGTyL1Δ454*mhil3AI* is identical to pGTyL1*mhil3AI* except for a 454-base 3' truncation, deleting the entire L1 3' UTR and the last 75 amino acids of L1.

The analysis of 13 histidine prototrophs is shown in Fig. 3a and b. Genomic DNA, digested with *HpaI*, was analysed by filter-blot hybridization. A *MAT* locus probe identifies the *MAT*-related sequences at *HML*, *HMR* and *MAT* (Fig. 3a). None of the clones retained the same sized band as the parent strain at the *MAT* locus, although the *HML* and *HMR* loci appear unaltered. Furthermore, the sizes of the restriction fragment(s) in place of the *MAT* band are heterogeneous, suggesting that each represents an independent rearrangement event. We rehybridized the filter with a *HIS3* probe (Fig. 3b). In each lane, the size of the *HIS3*-containing restriction fragment matches that of an altered *MAT* locus band, indicating that *HIS3*-containing cDNAs are linked to the *MAT* locus.

We measured the frequency of targeted *HIS3* insertions using a variety of plasmid constructs and yeast strains (Table 1). When the L1 RT was made non-functional by an active-site mutation (pGTyL1*RTmut**mhil3AI*), no histidine prototrophs were observed. Similarly, when a pG*mhil3AI* plasmid (lacking Ty1 or L1 sequences) was coexpressed with HO, no targeted insertions were observed. These results suggest that a functional RT protein is required to produce cDNA, and eliminates the possibility that spliced RNA is targeted to the DSB site and used as a template for repair by a cellular polymerase. When similar measurements were done in an isogenic *RAD52* strain, the overall frequency of histidine prototrophy increased 80-fold, but none of 18 independent clones examined had insertions at the *MAT* locus. This result indicates that cDNA insertions at the *MAT* locus site do not represent a major repair pathway in homologous recombination-proficient strains. Finally, deletion of the plasmid-based copy of the HO endonuclease gene eliminates targeting of *HIS3* insertions to the *MAT* locus, indicating that the *MAT* locus is not intrinsically a preferred site for targeted cDNA insertions; rather, targeting is a response to the presence of an HO-induced DSB.

We then investigated whether the targeting phenomenon was specific to L1. We induced equivalent *mhil3AI*-containing constructs

TABLE 1 Specificities of DNA targeting

Yeast strain	Frequency of histidine prototrophs among healed survivors	Proportions of <i>HIS3</i> insertions at <i>MAT</i> locus amongst histidine prototrophs	Derived frequency of <i>HIS3</i> insertions at <i>MAT</i> locus among healed survivors
<i>rad52</i> , HO+, pGTYL1Δ454 <i>mhis3Al</i>	6.8×10^{-6} (1.0)	0.909 (10/11)	6.18×10^{-6}
<i>rad52</i> , HO+, pGTYL1 _{RTmut} <i>mhis3Al</i> *	$< 2.2 \times 10^{-8}$	Not done	
<i>rad52</i> , HO+, pG <i>mhis3Al</i> †	6.0×10^{-9} (0.001)‡	Not detected (0/1)	
<i>rad52</i> , HO+, pGTYL1 <i>mhis3Al</i>	2.3×10^{-7} (0.03)	0.80 (4/5)	1.83×10^{-7}
<i>RAD52</i> , HO+, pGTYL1Δ454 <i>mhis3Al</i>	5.5×10^{-4} (80.9)	< 0.06 (0/18)	$< 3.67 \times 10^{-5}$
<i>rad52</i> , ho-, pGTYL1Δ454 <i>mhis3Al</i>	6.9×10^{-6} §	< 0.04 (0/22)	$< 3.83 \times 10^{-7}$
<i>rad52</i> , HO+, pGTY <i>mhis3Al</i> (wild type)	1.6×10^{-4} (23.5)	< 0.08 (0/12)	$< 1.33 \times 10^{-5}$
<i>rad52</i> , HO+, pGTY _{IN mutant} <i>mhis3Al</i>	1.9×10^{-5} (2.9)	0.56 (9/16)	1.1×10^{-5}
<i>rad52</i> , HO+, pGTYCRE1 <i>mhis3Al</i> *	5.7×10^{-7} (0.08)	0.667 (2/3)	3.79×10^{-7}

Either AGY49 or the *RAD52* isogenic YH50 strain were used, containing derivatives of the two plasmids shown in Fig. 2. Overall frequency of histidine prototrophy was estimated by first determining the ratio of double histidine and tryptophan prototrophs to single tryptophan prototrophs. This value was then corrected for the fraction of single and double prototrophs that were either sterile or *MATa*, as these represented cells that had been unequivocally cut by HO endonuclease and subsequently repaired²⁵. Frequencies are the mean of 18 independent inductions unless otherwise indicated. The proportion of *HIS3* insertions at the *MAT* locus was determined by genomic DNA blots and mating type testing. The frequency of *HIS3* insertions at the *MAT* locus among healed survivors was derived by multiplying the values obtained in the previous two columns. Frequency of non-*HIS3* insertions among healed survivors was determined by filter hybridization to be $\sim 1.4 \times 10^{-2}$. HO+ and ho- refer to the presence or absence of the functional CEN-based GAL-HO plasmid, pGHOT-GAL3. Relative frequencies in parentheses were calculated by dividing each mean frequency by that of pGTYL1Δ454*mhis3Al*.

* Frequency is the mean of 4 independent inductions.

† Frequency is the mean of 13 independent inductions.

‡ This represented a single histidine prototroph, probably resulting from the residual amount of expressed Ty in *spt3* strains.

§ Derived frequency values for this strain are not comparable with the others because no HO-induced cutting occurred, that is, all cells remained *MATa*.

in yeast, in which the L1 portion had been replaced by either wild type or a mutant version of Ty1, or by the site-specific non-LTR retrotransposon CRE1 (ref. 12) (Table 1). In the presence of a wild-type Ty1 element, we observed a high frequency of histidine prototrophs, but none, of 12 independent histidine prototrophs examined was associated with an insertion at the *MAT* locus. With an integrase-defective version of Ty1 (IN mutant), the overall frequency of histidine prototrophy decreased almost 10-fold, but a large fraction of prototrophs were linked to the *MAT* locus. The high level of histidine prototrophy seen with wild-type Ty1 probably results from normal Ty1 transposition (that is, integrase-mediated Ty1 chromosomal insertions), which is known to be *RAD52*-independent¹³. However, this high level of transposition also obscures detection of potential targeted cDNA insertions. Finally, using the CRE1 ORF as a third source of RT (ref. 14, and S.C.T. and A.G., unpublished observations), two of three independent histidine prototrophs show linkage to the *MAT* locus. The factors common to experiments yielding high levels of targeted cDNA insertions were a functional source of RT within Ty1 VLPs, a persistent DSB at *MAT*, and the absence of an alternative pathway for cDNA insertion.

The structure of representative insertion events generated by three different RT sources is shown in Fig. 3c. In the simplest cases (insertions 26 and 33), the insertions consist of spliced *HIS3* genes flanked both upstream and downstream by contiguous sequences present on the plasmid. Insertion 7 shows evidence of RT template switching during cDNA synthesis, with the *HIS3* gene linked by means of a non-templated poly(A) stretch to the 5' end of the Ty1-H3 transcript (that is, the 5' U3-R junction). Analysis of junctional sequences from many insertions (Fig. 3c and data not shown) demonstrates several phenomena. First, the *HIS3* gene is inserted in either orientation relative to *MAT*, suggesting the absence of directionality for the repair process. Second, in most cases the junctions with *MAT* contain 1–4 bases of overlapping 'microhomology' between the *MAT* sequence and the expected cDNA sequence. Microhomology regions are a signature of both mammalian (reviewed in ref. 15) and yeast (reviewed in ref. 11) non-homologous recombination events. Third, in all cases, a variable number of bases are deleted from both sides of the HO

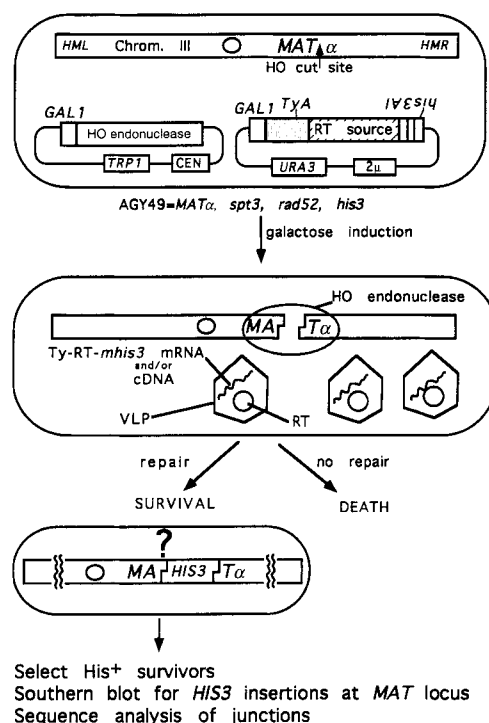
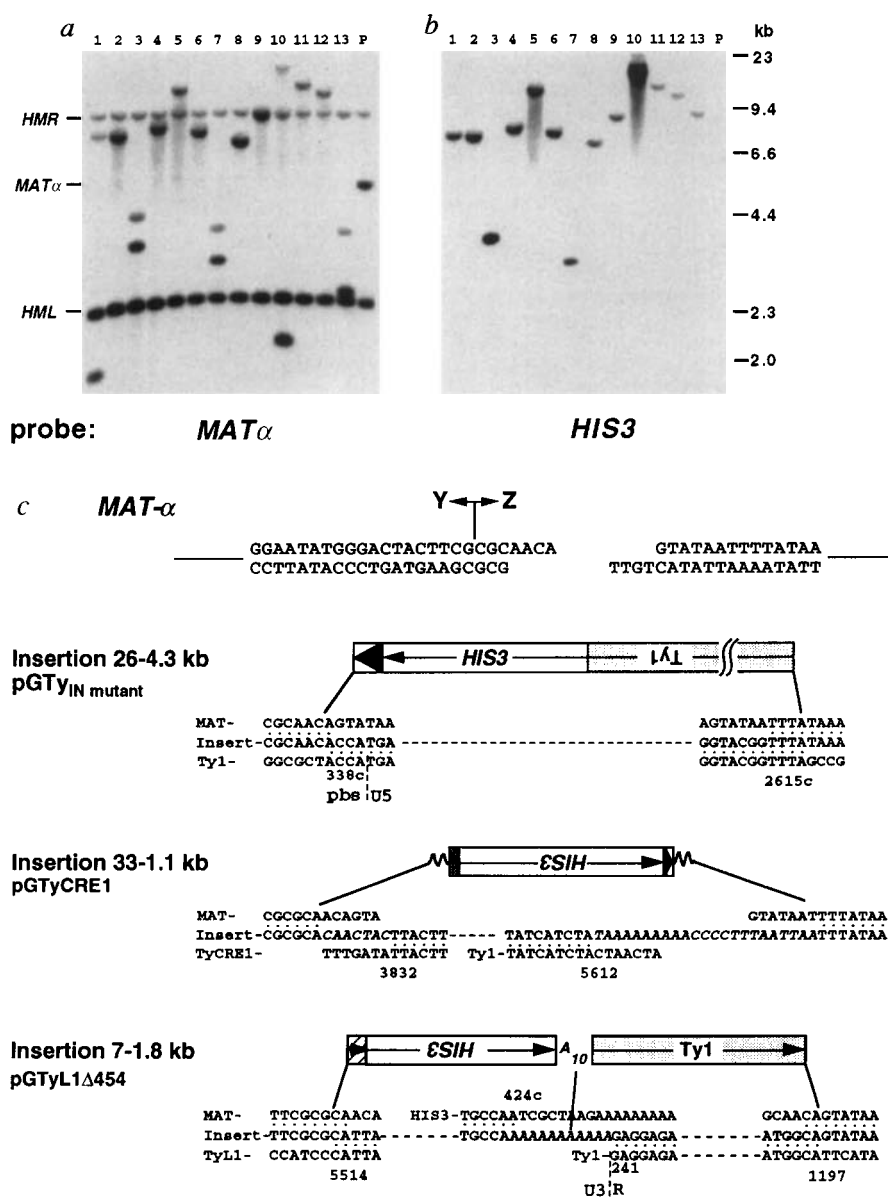


FIG. 2 Scheme for trapping RT-mediated DSB repair events. Strain AGY49 lacks endogenous Ty1 expression (*spt3*) as well as homologous recombination (*rad52*). Two plasmids are in this strain. pGHOT-GAL3 contains a GAL1-promoted HO endonuclease gene. The second consists of a derivative of a GAL1-promoted Ty1 element (in which most of the TyB ORF may be replaced by a heterologous source of RT) with a downstream *mhis3Al* marker gene. On galactose induction, the HO endonuclease makes a unique DSB at the *MAT* locus, which, if not repaired, is lethal. Simultaneously Ty1 VLPs are formed containing a source of RT and a spliced *HIS3* mRNA. Survivors of the induction period that are histidine prototrophs have undergone an RT-mediated insertion event, and can be isolated and analysed further.

FIG. 3 Analysis of *HIS3* chromosomal insertions. *a* and *b*, Genomic DNA from 13 independent *His*⁺ events derived from pGTYL1Δ454*mhis3AI* (lanes 1–13) and the parental strain (lane P) were digested with *Hpa*I, which cuts outside the *MAT*, *HML*, and *HMR* loci, and analysed by filter hybridization using *MAT* (*a*) or *HIS3* (*b*) probes. Extra bands observed in *a* for insertions 3, 7 and 13, are due to a *Hpa*I site in the Ty1 portion of the insert. The endogenous *HIS3* gene is deleted in the parental strain. *c*, The sequence of the HO endonuclease cut site in the *MAT*_α locus is followed by the structure and junctional sequences of three insertion events mapped to the *MAT* locus. Junctional overlap sequences are marked by dots. Italicized bases represent non-templated sequences. The orientation of each gene is indicated by the orientation of its acronym, and the orientation of the mRNA undergoing reverse transcription is indicated by an arrow. Reference nucleotide coordinates for each gene are shown below the junctions. The presence of four bases of the Ty1 primer binding site (pbs) at the left junction of insertion 26 suggests that a portion of the initiator methionine tRNA (which primes minus-strand synthesis) took part in the recombinational event. The way lines flanking *HIS3* in insertion 33 indicate non-templated bases at the junction. The 23 bases at the right junction were not found in a search of the complete *Saccharomyces cerevisiae* genome. The central junction between the U3/R segment of Ty1, which represents the 5' end its transcript, and the 10 non-templated A residues of insertion 7, suggests that the L1 RT reached the end of the Ty1 transcript and transferred to the polyadenylated tail of an antisense *HIS3*-containing mRNA.



cut site, suggesting that the timing of repair is variable relative to the digestion of the ends. Finally, in several events (for example, insertion 33), the junctional sequences contain multiple bases of unknown origin. A plausible explanation is that these bases represent RT-mediated non-templated base additions at the cut site, as has been observed previously for the L1 and R2Bm RTs^{5,16}.

cDNA-derived sequences found in mammalian genomes have been postulated to represent the 'fossil record' of RT-mediated insertions at sites of chromosomal breaks⁹. Our results and those of Moore and Haber¹⁷ indicate that a repair mechanism linking non-homologous recombination and retrotransposition operates in eukaryotic organisms, providing evidence for a new, biologically relevant mechanism for chromosome repair in cells containing RT. Although our data demonstrate that cDNAs are targeted to a DSB site, they do not indicate the mechanism by which repair is accomplished. A simple model would involve RT-directed synthesis of a cDNA and migration of the cDNA to the break site, possibly in conjunction with Ty1 VLPs. The cDNA could then serve as a template for polymerization by the plasmid-encoded RT or a cellular DNA polymerase, with priming at the break site. The presence, in some insertions, of non-templated junctional sequences supports a direct role for RT in some cut site-primed

DNA synthesis events. Such events might share mechanistic features with other recently identified RT-mediated phenomena, such as the concerted insertion and reverse transcription of the R2Bm non-LTR retrotransposon¹⁸, telomere formation in *Drosophila*¹⁹, and yeast group II intron homing²⁰. □

Methods

Strains and plasmids. Assays were performed in yeast strains YH50 (*MAT*_α *ho* *spt3-202 his3Δ200 ura3-167 trpΔ1 leu2-1*)³ or AGY49, which was derived from YH50 by one-step gene replacement of the *RAD52* gene with *rad52::LEU2*. All plasmids were transformed into yeast by a lithium acetate method²¹. The plasmid p*Gmhis3AI* is referred to as pSM50 in ref. 3. pGTYL1*mhis3AI* was constructed by ligating the 990-bp *mhis3AI*-containing *Clal* fragment from BJC267 (ref. 22) with plasmid pSM2 (ref. 8) at its unique *Clal* site 3' to L1 ORF2, and selecting transformants with an antisense orientation of the *HIS3* ORF relative to Ty1-L1 ORF2 (BK59). BK59 was digested with *Bsp*EI, and the larger restriction fragment was self-ligated. The resulting plasmid, pGTYL1*mhis3AI*, lacks both the Ty1 3' LTR and a *TRP1* marker present in BK59. pGTYL1_{RTmut}*mhis3AI* was constructed identically to pGTYL1*mhis3AI* except that pSM8 (ref. 8) replaced pSM2 in the cloning. pGTYL1*mhis3AI*Δ454 was constructed by partially digesting BK59 with *Clal*, end-filling and then self-ligating. BK274, selected because the downstream *Clal* site had been lost, was digested with *Clal*, and the linear fragment was subjected to timed bidirectional nucleolytic digestion (Erase-a-Base;

Promega). The resulting deleted fragments were blunt-ended and cut with *Afl*III. A fragment of ~1,400 bp was ligated to BK274, which had previously been digested with *Cl*al, end-filled, and then redigested with *Afl*III. The junction of the deletion was determined by direct sequencing of the resulting plasmid. Subsequently the plasmid was digested with *Bsp*EI, as with BK59, and religated. pGHOT-GAL3 is a *TRP1* marked centromere-based plasmid containing a *GAL1*-promoted HO endonuclease gene. The ho⁻ derivative of pGHOT-GAL3 was constructed by deleting the small *Bam*HI fragment, thereby removing the 5' half of HO endonuclease. pGTY^(Nmutant)*mhis3*ΔI was constructed by replacing the 2.8-kb *Bst*EII-*Afl*III fragment from pGTY*mhis3*ΔI with the equivalent fragment from pGM119, which contains an IN-inactivating linker insertion²³. pGTYCRE¹⁴*mhis3*ΔI was constructed by first adding the 3' LTR region of Ty1-H3 (nucleotides 5585–5918) to the unique *Hind*III site downstream of CRE1 in pGTYCRE¹⁴ after PCR amplification of pJEF724 (ref. 24) using primers RAG 153 (5'-CGCTAAAGCT-TACGCGTCGCTGGTACCATCGATAGATCTATTACATTATGG-3') and RAG 154 (5'-CGCTAAAGCTTACAAATTGATAAGCAATG-3'). The 990-bp *mhis3*ΔI fragment from BJC267 was then cloned into the resulting unique *Cl*al site, as for the pGTYL1*mhis3*ΔI construction.

Genetic assays. Patch assays were performed by replica plating yeast patches onto synthetic complete (SC) media plates lacking uracil, supplemented with 2% galactose, and incubating for 4 days at 22 °C (induction), before replica plating to synthetic complete media lacking histidine, supplemented with 2% glucose, and growing for 3 days at 30 °C. For isolation of insertion events and determination of frequencies of histidine prototrophy, single colonies growing on selective media plates were inoculated into 5 ml SC media lacking tryptophan and uracil and supplemented with raffinose to 1%, grown at 30 °C until reaching an absorbance of 0.5 at 600 nm and then induced at 30 °C by the addition of galactose to 2%. After 24 h induction, appropriate dilutions of cells were plated on YPD, SC minus tryptophan plus glucose, and SC minus histidine and tryptophan plus glucose plates, and incubated for 4 days at 30 °C before counting. Individual colonies were isolated and tested for mating types.

DNA manipulations. Independent His⁻ papillae were grown to saturation, DNA was extracted, digested with *Hpa*I, separated through 0.5% agarose, transferred to GeneScreen plus (DuPont) nylon membranes, and hybridized at high stringency with ³²P-labelled random primed DNA probes. Filters were stripped between hybridizations. The *MAT* probe consisted of a 1,358-bp PCR product, spanning nucleotides 394–1752 of the *MAT* region. The *HIS3* probe consisted of the 990-bp *Cl*al fragment from BJC267 (ref. 22). Genomic DNA from selected insertion strains was subjected to PCR amplification using combinations of primers located in the regions flanking the *MAT* locus and/or within *HIS3*, *L1* or *Ty1* genes. PCR products were sequenced either directly or after cloning into a TA (Invitrogen) vector. Reference sequences: *L1*.2, Genbank accession no. M80343; *HIS3*, Genbank accession no. X03245; *Ty1*-H3, Genbank accession no. M18706; *MAT*Δ, Genbank accession no. J01332; CRE1, Genbank accession M33009.

Received 22 May; accepted 26 July 1996.

- Boeke, J. D., Garfinkel, D. J., Styles, C. A. & Fink, G. R. *Cell* **40**, 491–500 (1985).
- Derr, L. K. & Strathern, J. N. *Nature* **361**, 170–173 (1993).
- Dombroski, B. A. et al. *Mol. Cell. Biol.* **14**, 4485–4492 (1994).
- Nissley, D. V., Garfinkel, D. J. & Strathern, J. N. *Nature* **380**, 30 (1996).
- Zou, S., Wright, D. A. & Voytas, D. F. *Proc. Natl Acad. Sci. USA* **92**, 920–924 (1995).
- Curcio, M. J. & Garfinkel, D. J. *Proc. Natl Acad. Sci. USA* **88**, 936–940 (1991).
- Derr, L. K., Strathern, J. N. & Garfinkel, D. J. *Cell* **67**, 355–364 (1991).
- Mathias, S. L., Scott, A. F., Kazanian, H. H. Jr, Boeke, J. D. & Gabriel, A. *Science* **254**, 1808–1810 (1991).
- Hutchison, C. A., Hardies, S. C., Loeb, D. D., Shehee, W. R. & Edgell, M. H. in *Mobile DNA* (eds Berg, D. E. & Howe, M. M.) 593–617 (American Society of Microbiology, Washington DC, 1989).
- Herskowitz, I., Rine, J. & Strathern, J. N. in *The Molecular and Cellular Biology of the Yeast Saccharomyces* (eds Jones, E. W., Pringle, J. R. & Broach, J. R.) 583–656 (Cold Spring Harbor Laboratory Press, NY, 1992).
- Haber, J. E. *Bioessays* **17**, 609–620 (1995).
- Gabriel, A. et al. *Mol. Cell. Biol.* **10**, 615–624 (1990).
- Sharon, G., Burkett, T. J. & Garfinkel, D. J. *Mol. Cell. Biol.* **14**, 6540–6551 (1994).
- Gabriel, A. & Boeke, J. D. *Proc. Natl Acad. Sci. USA* **88**, 9794–9798 (1991).
- Roth, D. & Wilson, J. in *Genetic Recombination* (eds Kucherlapati, R. & Smith, G. R.) 621–653 (American Society for Microbiology, Washington DC, 1988).
- George, J. A., Burke, W. D. & Eickbush, T. H. *Genetics* **142**, 853–863 (1996).
- Moore, J. K. & Haber, J. E. *Nature* **383**, 644–646 (1996).
- Luan, D. D., Korman, M. H., Jakubczak, J. L. & Eickbush, T. H. *Cell* **72**, 595–605 (1993).
- Sheen, F. & Levis, R. W. *Proc. Natl Acad. Sci. USA* **91**, 12510–12514 (1994).
- Zimmerly, S., Guo, H., Perlman, P. S. & Lambowitz, A. M. *Cell* **82**, 545–554 (1995).
- Gietz, D., St Jean, A., Woods, R. A. & Schiestl, R. H. *Nucleic Acids Res.* **20**, 1425 (1992).
- Curcio, M. J. & Garfinkel, D. J. *Genetics* **136**, 1245–1259 (1994).
- Monokian, G. M., Brateman, L. T. & Boeke, J. D. *Gene* **139**, 9–18 (1994).
- Boeke, J. D., Xu, H. & Fink, G. R. *Science* **239**, 280–282 (1988).
- Moore, J. K. & Haber, J. E. *Mol. Cell. Biol.* **16**, 2164–2173 (1996).

ACKNOWLEDGEMENTS. We thank J. Boeke, J. Curcio, M. Fasullo and L. Symington for providing plasmids; H.-Y. Liu and R. Wilson for technical help; members of our laboratory and J. Boeke, M. Gartenberg, J. Haber, D. Livingston, D. Norris and S. Sandmeyer for discussion; and M. Driscoll, M. Gartenberg and E. Mules for reading the manuscript. This work was supported in part from grants from the NIH, NSF, the Charles and Johanna Busch Endowment and the Lucille P. Markey Charitable Trust. A.G. is a Lucille P. Markey Scholar.

CORRESPONDENCE and requests for materials should be addressed to A.G. (e-mail: gabriel@mbcl.rutgers.edu).

Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks

J. Kent Moore & James E. Haber

Rosenstiel Center and Department of Biology, Brandeis University, Waltham, Massachusetts 02254–9110, USA

NON-HOMOLOGOUS repair of broken chromosomes in *Saccharomyces cerevisiae* can be studied at a defined location by expressing the site-specific HO endonuclease that cuts the mating-type (*MAT*) locus. When homologous recombination is prevented, most double-strand breaks are repaired by non-homologous end-joinings similar to those observed in mammalian cells. About 1% of non-homologous repair events were exceptional, having 'captured' approximately 100 base pairs of DNA within the HO cleavage site. In each case, the insertion came from yeast's retrotransposon Ty1 element. Four of the five contained the R-U5 region, which is the first part of Ty1 messenger RNA to be converted to complementary DNA. The capture of cDNA fragments at the sites of double-strand breaks may account for the way that pseudogenes and long and short interspersed sequences (LINES and SINES) have been inserted at many locations in the mammalian genome.

Double-strand breaks in eukaryotes are repaired by several pathways of homologous recombination (reviewed in ref. 1) or by non-homologous end-joining, where the ends of the junction share 0–5 base pairs (bp) of homology^{2–5}. In *S. cerevisiae*, the repair of double-strand breaks at defined chromosomal locations has been studied by expressing the site-specific HO endonuclease to cleave the *MAT* locus^{1,4–7}. To examine non-homologous repair in wild-type cells, we created a haploid strain in which both the *HML* and *HMR* donors that normally repair (and switch) *MAT* have been deleted⁸, and so homologous recombination is eliminated. The *HO* gene is under the control of a galactose-inducible promoter⁶. When cells are plated on galactose-containing plates, about 0.2% survive, both in wild-type and *rad52* strains. Survival is 70-fold lower in strains deleted for *rad50*, *xrs2* or *mre11* (ref. 5).

Nearly all survivors have mutations in the *MAT*Δ1 gene which traverses the HO cutting site; they are identified as sterile (*mat*Δ1) cells⁵. The nature of non-homologous repair events was determined by sequencing regions around the cutting site that were amplified by the polymerase chain reaction (PCR). There are two predominant types of repair^{4,5}. In both wild-type and *rad52* strains nearly 80% contain 2- or 3-bp insertions, and most of the remainder have small deletions (Fig. 1a). The small insertions are nearly eliminated in *rad50*, *xrs2* and *mre11* strains. However, three survivors were exceptional and yielded PCR products ~100 bp larger than normal (Table 1). Their DNA sequences showed that each contained a part of yeast's retrotransposon, Ty1.

TABLE 1 Frequency of survival by non-homologous end-joining

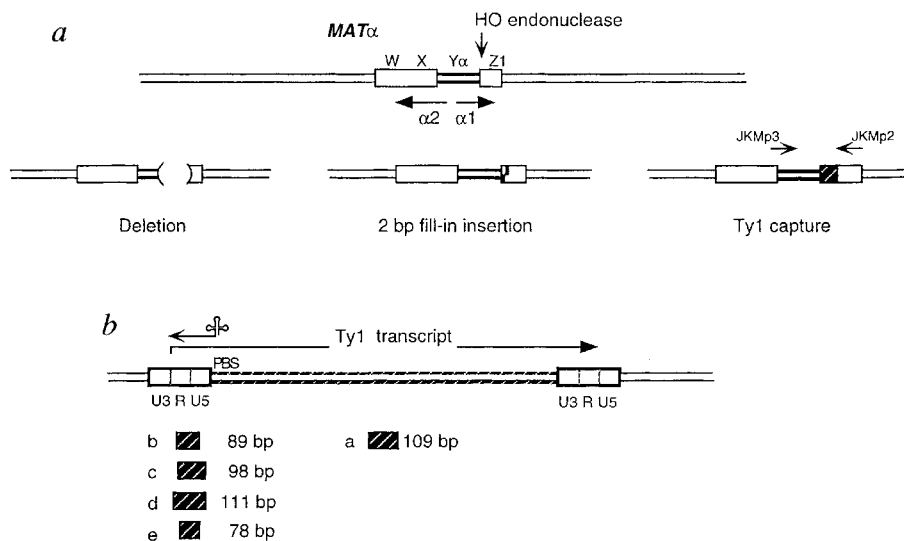
Strain	Survival frequency*	Total analysed	Large insertions
wild type	2.2×10^{-3}	85	1
<i>rad52</i>	2.0×10^{-3}	39	1
<i>rad50</i>	3.0×10^{-5}	72	0
<i>xrs2</i>	2.2×10^{-5}	24	0
<i>mre11</i>	3.0×10^{-5}	72	1
wild type + pJEF1105†	4.0×10^{-3}	164	2

Cells that survived HO cleavage were analysed as described in Fig. 1.

* Except for the wild-type strain with pJEF1105, survival data are from ref. 5.

† pJEF1105 is a centromeric plasmid carrying Ty1 under control of a galactose-inducible promoter. PCR analysis detected repair events in which the size of the cutting-site region increased by >50 bp.

FIG. 1 Non-homologous end-joining of a broken chromosome. *a*, The general structure of the *MAT α* locus is shown. HO endonuclease creates a double-strand break (DSB) near the Y/Z border that must be repaired for the cell to survive. Two major pathways of DNA end-joining have been defined⁵. The overlapping ends of the DSB can be misaligned and filled in to produce a small (2 or 3 bp) insertion that ablates the HO cleavage site. Alternatively, DNA ends are degraded, to create a deletion, presumably by base pairing of single-stranded regions generated by 5' to 3' degradation of the cut DNA ends. A third pathway involves capture of Ty1 DNA into the DSB site at *MAT α* . All events create a *mat α 1* (sterile) mutation. *b*, General structure of the retrotransposon Ty1, including the three subdomains of the LTR δ sequence and the adjacent primer binding site (PBS). The rightward mRNA transcript is indicated together with the first step in cDNA synthesis in which a tRNA primes leftward cDNA synthesis, beginning at the PBS¹². The approximate locations of five segments of Ty1 DNA that are inserted into *MAT α* are shown.



The insertion in the wild-type derivative contained a 109-bp fragment of DNA from the middle of the retrotransposon Ty1 variant, Ty-pY109 (ref. 8) (a in Fig. 1b; Fig. 2). The exceptional *rad52* derivative carried an 89-bp insertion that is 99% identical to the long terminal repeat (LTR) δ element that is found at the ends of Ty1 elements and also as 'solo' δ sequences dispersed throughout the yeast genome^{9,10} (b in Fig. 1b; Fig. 2). The 5' end of this 89-bp sequence is nearly coincident with the transcription start site of Ty1 mRNA¹¹. An adenine residue at the junction of the 3' end of the Ty1 segment with *MAT* may reflect a base-pair variant in Ty1 sequences. In the *mre11* strain, there was a 98-bp insertion (c in Fig. 1b; Fig. 2) in which the 5' end begins 4 bp from the transcription start site of Ty1 mRNA and extends past the end of U5 into the 10-bp primer binding site, where an initiator methionine transfer RNA binds to initiate cDNA synthesis¹².

We then investigated whether simultaneous overexpression of Ty1-encoded reverse transcriptase and other Ty1 proteins affected the frequency of Ty1 capture. Wild-type cells that lack *HML* and *HMR* and carrying *GAL::HO* plasmid pFH800 (ref. 13) were transformed with plasmid pJEF1105 (ref. 14), which expresses Ty1 mRNA under control of the *GAL1* promoter. The overall survival of these cells was similar to those carrying only a *GAL::HO* plasmid (Table 1). Among 164 sterile (*mat α 1*) survivors examined by PCR, two had Ty1 insertions. This frequency is the same as that found in the absence of pJEF1105. Both of these Ty1 segments include part of the R and U5 segments of the LTR region, and one extends into the U3 region (d and e in Fig. 1b; Fig. 2). The additional four base pairs in the U5 region of event E (CAAA) have been reported for a solo δ insertion in the *CYC7* gene¹⁵.

The junctions between Ty1 sequences and the ends of the HO-cut fragments resemble the junction of other non-homologous end-joining events^{2,4,16}. Of the 10 junctions, 4 had 2–3 bp of overlapping homology, 3 had 1 bp and 3 had none; 5 of the junctions had also removed between 1 and 16 bp of the HO-cut *MAT* DNA; in the others, the 3' ends created by HO were apparently intact. All four events that captured the R-U5 region have junctions not in U5 but within the primer binding site region.

In wild-type and *rad52* strains, the frequency of acquiring Ty1 DNA is approximately 3×10^{-5} of all HO-cut cells (2 of 124 events among 2×10^{-3} survivors). Capture of Ty1 DNA is apparently reduced in *mre11*, *rad50* and *xrs2* mutant strains. Although the overall frequency of capture was 1 out of 166 independent survivors among the three phenotypically similar mutant strains (compared to 2 of 124 in wild-type and *rad52* cells), the overall survival rate among HO-cut cells was reduced 70-fold. If Ty1 captures had been insensitive to mutations in *rad50*, *xrs2* or *mre11*,

their proportion among surviving colonies should have been much greater.

That all five examples of captured DNA contain part of Ty1 is clearly not expected by chance, as there are fewer than 50 copies of the 6.7-kb Ty1 DNA in a genome of 15,000 kb. It seems likely that their capture is a consequence of Ty1 retrotransposition. Some Ty1 cDNA may be released from the virion, where its replication normally occurs, and may be captured by non-homologous recombination at sites of pre-existing double-strand breaks. Further evidence supporting this idea comes from the finding¹⁷ that integration of Ty1-mediated cDNA containing the selectable *HIS3* gene¹⁸ can also occur preferentially at the site of an HO-induced double-strand break. Capture of large *HIS3*-containing cDNAs was much less frequent than for the segments of ~100 bp that we obtained (3.7×10^{-6} versus 2×10^{-2}), although the former reflects a much more stringent selection and does not include other insertions that do not include an intact *HIS3* gene.

In our non-selective system, only small (~100 bp) segments of Ty1 DNA were captured. Four of the five events involve sequences that participate in the initial steps of converting Ty1 mRNA to cDNA, in which a tRNA primes cDNA synthesis of the R and U5 region of the LTR from the primer binding site¹². Indeed, when Ty1 virions are induced in yeast, a single-stranded fragment of cDNA of ~95 bp was recovered from yeast cells¹⁹. It seems important that all of these four events contain a few nucleotides of the primer binding site sequence. It is possible that capture of this region of Ty1 involves the integration or replication of a cDNA segment in which the 5' end is part of the primer tRNA. Reverse splicing of RNA into a DNA site has recently been described for group II introns²⁰.

Our results and those of another study¹⁷ provide complementary evidence about the process of inserting cDNA fragments into the genome at sites of DNA damage. This process can account for the insertion of large cDNA fragments, such as pseudogenes, into the genome¹⁷. Our study shows that such cDNA capture at the site of double-strand breaks can occur without inducing a high level of retrotransposon mRNA. Indeed, there was no significant increase in Ty1 capture when HO and Ty1 were expressed simultaneously. Thus, although inducible Ty1 expression is clearly necessary to follow a specifically marked Ty1-associated cDNA, at least the initial intermediates of endogenous Ty1 cDNA synthesis are present in significant quantities in yeast to participate in the repair events we have observed. We cannot exclude the possibility that these derive from mutant Ty1 elements incapable of carrying out complete retrotransposition, although at least three different Ty1 sequences have been used.

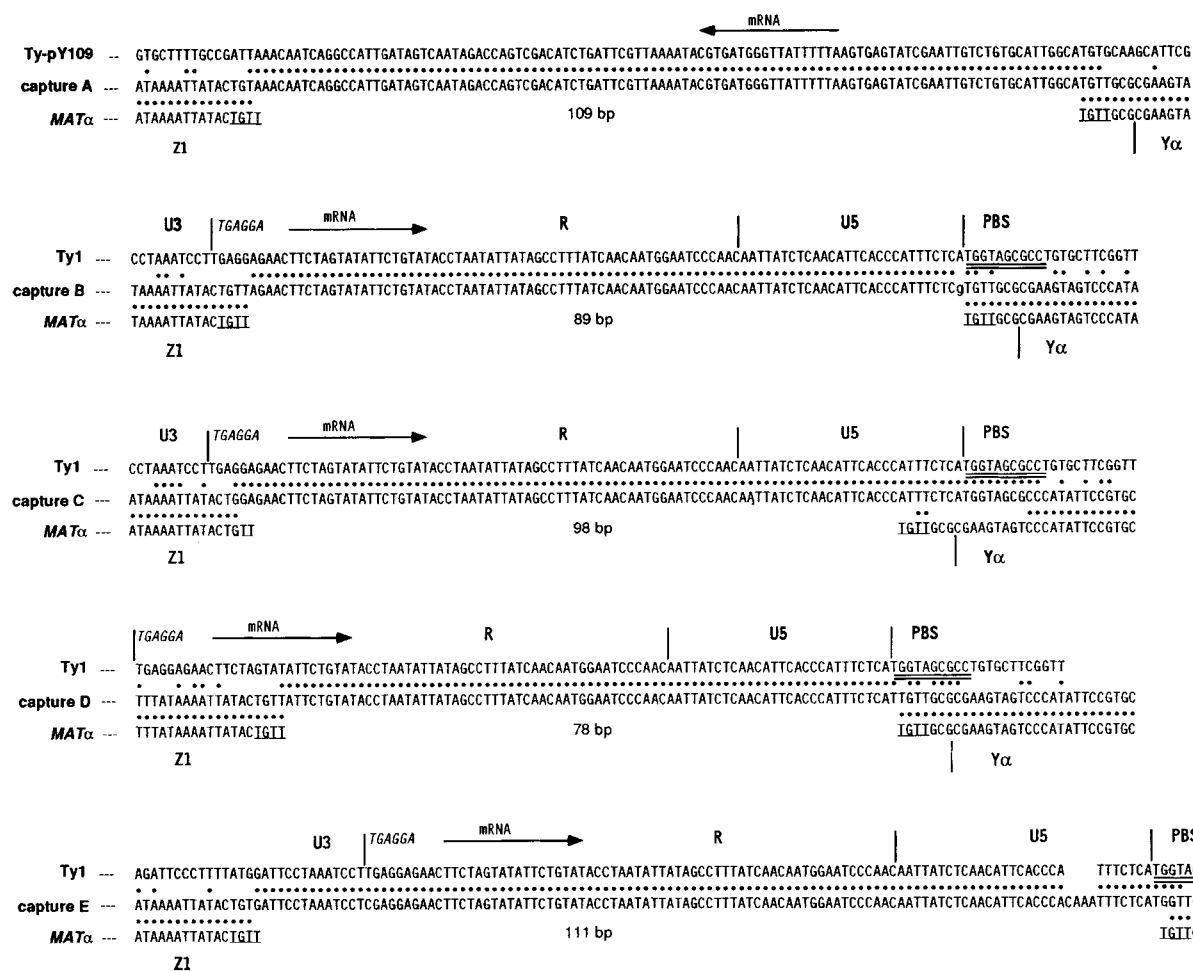


FIG. 2 DNA sequence of the captured Ty1 segments in the MAT locus. The DNA sequence of each inserted segment is shown in alignment with MAT α and Ty1 DNA sequences. Dots indicate identical DNA sequences shared with one or the other sequence. The location of the 5' end of the Ty1 mRNA

transcript is shown in the four cases where Ty1 DNA is derived from the U3-R-U5 regions of the LTR and adjacent PBS sequences (double underlined). The 4-bp overhanging 3' ends of the HO cut at MAT are underlined.

This mechanism may be important in mammalian cells as well as in yeast. The mammalian genome contains a large number of dispersed copies of SINES, LINES and pseudogenes that are thought to have been created by reverse transcription and subsequent integration^{21,22}. That sequences such as Alu can promote their own transcription shows how they provide a substrate for reverse transcription, but does not explain how the cDNA is inserted at a new genomic site. We suggest that the capture of these cDNAs occurs at sites of random or programmed DNA breakage.

Note added in proof: The capture of the R-U5 region of Ty1 DNA into the genome has apparently occurred several times in yeast's evolutionary history. A survey of the entire yeast genome (<http://genome-www.stanford.edu/Saccharomyces/>) reveals five locations that contain the R-U5 regions of the LTR without containing U3 or other portions of Ty1. □

Methods

All strains were derivatives of JKM110 (ref. 5), which has the genotype *hoΔ MATα hmlΔ::ADE1 hmrΔ::ADE1 leu2-3,112 ura3-52 lys5 ade1* and carries the centromeric URA3-marked plasmid pGAL-HO, bearing a galactose-inducible HO gene⁵. This strain is unable to repair a double-strand break (DSB) by homologous recombination. The creation of derivatives deleted of *RAD50*, *RAD52*, *XRS2* and *MRE11* are also described in ref. 5. Cultures of JKM110 or its derivatives were plated on agar plates containing 2% galactose plus synthetic medium lacking uracil. Colonies that arose were resistant to HO-induced cleavage of the HO cleavage site in MAT α . Non-mating (sterile) derivatives were presumed to have altered the HO cleavage site in the MAT α 1 gene⁵. Two oligonucleotide primers, JKMp2 and JKMp3 (Fig. 1a), were used to amplify the MAT α region⁵. JKMp2 was

also used as a DNA sequencing primer. PCR-amplified DNA was run on 1.5% seakem agarose gels (FMC) to determine if there had been a deletion or insertion of at least 50 bp of DNA, the limit of resolution of this method using agarose. PCR fragments were purified for DNA sequence analysis using a Prep-a-gene DNA purification matrix (Bio-Rad).

Received 20 May; accepted 29 July 1996.

1. Haber, J. E. *BioEssays* **17**, 609–620 (1995).
2. Roth, D. & Wilson, J. H. in *Genetic Recombination* (eds Kucherlapati, R. & Smith, G. R.) 621–653 (American Society for Microbiology, Washington DC, 1988).
3. Pfeiffer, P., Throde, S., Hanke, J. & Vielmetter, W. *Mol. Cell. Biol.* **14**, 888–895 (1994).
4. Kramer, K. M., Brock, J. A., Bloom, K., Moore, J. K. & Haber, J. E. *Mol. Cell. Biol.* **14**, 1293–1301 (1994).
5. Moore, J. K. & Haber, J. E. *Mol. Cell. Biol.* **16**, 2164–2173 (1996).
6. Jensen, R. & Herskowitz, I. *Cold Spring Harb. Symp. Quant. Biol.* **49**, 97–104 (1984).
7. Connolly, B., White, C. I. & Haber, J. E. *Mol. Cell. Biol.* **8**, 2342–2349 (1988).
8. Hauber, J., Nelbock, J. H. & Feldmann, H. *Nucleic Acids Res.* **13**, 2745–2758 (1985).
9. Voytas, D. F. & Boeke, J. D. *Trends Genet.* **9**, 421–427 (1993).
10. Clare, J. & Farabaugh, P. *Proc. Natl Acad. Sci. USA* **82**, 2829–2833 (1985).
11. Boeke, J. D. & Corces, V. G. *Annu. Rev. Microbiol.* **43**, 403–434 (1989).
12. Chapman, K. B., Bystrom, A. S. & Boeke, J. D. *Proc. Natl Acad. Sci. USA* **89**, 3236–3240 (1992).
13. Nickoloff, J. A., Singer, J. D., Hoekstra, M. F. & Heffron, F. J. *Mol. Biol.* **207**, 527–541 (1989).
14. Boeke, J. D., Xu, H. & Fink, G. R. *Science* **239**, 280–282 (1988).
15. Errede, B., Company, M. & Swanson, R. *Mol. Cell. Biol.* **6**, 1334–1338 (1986).
16. Mézard, C. & Nicolas, A. *Mol. Cell. Biol.* **14**, 1278–1292 (1994).
17. Teng, S.-C., Kim, B. & Gabriel, A. *Nature* **383**, 641–644 (1986).
18. Derr, L. K., Strathern, J. N. & Garfinkel, D. J. *Cell* **67**, 355–364 (1991).
19. Müller, F., Laufer, W., Pott, U. & Ciriacy, M. *Mol. Gen. Genet.* **226**, 145–153 (1991).
20. Yang, J., Zimmerly, S., Perlman, P. S. & Lambowitz, A. M. *Nature* **381**, 332–335 (1996).
21. Britten, R. J., Stout, D. B. & Davidson, E. H. *Proc. Natl Acad. Sci. USA* **85**, 4770–4774 (1988).
22. Shen, M. R., Batzer, M. A. & Deininger, P. L. *J. Mol. Evol.* **33**, 311–320 (1991).

ACKNOWLEDGEMENTS. We thank A. Gabriel, M. Moore and members of our laboratory for suggestions. This work was supported by a grant from the US Department of Energy.

CORRESPONDENCE and requests for materials should be addressed to J.E.H. (e-mail: haber@hydra.rose.brandeis.edu).

Immunological toolbox

This section is an inventory of monoclonal antibodies and other new tools available for immunology, including a DAB system for immunoblotting, an HPLC system that monitors IgG glycosylation and the v-Ha-ras(TG.AC) transgenic mouse.

THE 'plus' version of the OptiMax immunostainer from BioGenex includes new features that provide part of a completely reliable, automated system of immunohistochemical staining. New features include increased reagent capacity, flexibility, automated cleaning, dual pumps for on-line fluids, and a printer interface for easier quality control and record-keeping. The system includes sen-



The BioGenex OptiMax automated immunostainer.

sitive pre-diluted, pre-optimized antibodies and detection reagents, an antigen retrieval pretreatment technology and computer-controlled automation of the traditional horizontal staining technique. According to the manufacturer, this instrument is easy to operate, can be set up in as little as 10 min, has removable slide and reagent racks for refrigerated storage, as well as random-access operation. This latter feature allows up to 33 protocols with varying reagents and incubation times to be run in a single automated cycle. The instrument is stated to produce quality results on up to 40 slides in under two hours. It is capable of performing multiple antibody titrations and can handle multiple links, labels and chromogens in a single staining run. The instrument's flexible system software allows development and storage of up to 250 custom procedures for both clinical and research applications. Upgrade options include a colour monitor, a Pentium CPU, Windows95 software and bar-code labelling.

Reader Enquiry No. 100

The new BX40CY microscope from Olympus incorporates features designed to simplify cytology screening work. The

system has a new frame and stage, $\times 10$ objective, a slide marker and a range of slide holders that retain the ergonomic design and UIS optics of the BX series. The BX40CY is built around a new frame, which has been designed for the repetitive inspection of the large number of specimens encountered in cytology screening. The stage support is 30 mm lower than other BX microscopes and the 172-mm stage height provides easier placement of specimens and reduces arm fatigue. Coaxial stage, control knobs are located in a position that allows the user's arms to rest on the table surface. The fixed five-position nosepiece is also lower, 221 mm above the table surface, to simplify the change of objectives. A new $\times 10$ objective has been developed with a built-in neutral density filter to equalize light intensity within the field of view. This allows switching between $\times 10$ and $\times 40$

magnification and back again at the same light intensity setting.

Reader Enquiry No. 101

DAB (3, 3'-diaminobenzidine tetrahydrochloride) substrate systems have been designed for immunohistological and immunoblotting procedures and are now available from Sigma. The DAB liquid enzyme substrate systems contain pre-measured and pre-mixed volumes of reagents for the visual detection of enzyme-linked antibodies in immunoassays. The weighing, waiting, filtering

and waste often associated with traditional preparation techniques is said to be eliminated; the user simply adds the liquid chromagen to the liquid buffer and mixes. The systems are packaged in quantities convenient for staining either individual slides or batches of slides, as well as immunoblotting. The dropper system is recommended for immunohistological staining, as the dropper bottles make it simple to apply the substrate to tissue sections fixed on slides. The standard system is suitable when the preparation of a variable quantity of DAB substrate is required, such as in immunoblotting or in the preparation of numerous immunohistological slides. The substrate system consists of a 10 $\times$ stock solution of DAB chromagen and a corresponding volume of DAB liquid buffer/peroxide solution. The substrate dropper system contains 10 $\times$ (stock of DAB chromagen and ten dropper bottles of DAB liquid buffer/peroxide solution. Each is certified stable for at least one year and is supplied with a comprehensive datasheet containing directions for use, references and a troubleshooting guide.

Reader Enquiry No. 102

A new HPLC system has been developed by Oxford GlycoSystems for simultaneous, quantitative separation of highly heterogeneous mixtures of fluorescently labelled neutral and acidic glycans. Subpicomolar amounts of the glycans can be resolved to provide a detailed glycosylation profile or fingerprint. Profile interpretation is made easier by the predictability of glycan elution. The normal-phase GlycoSep N HPLC column can be used to monitor changes in IgG glycosylation, both in

In order to simplify and broaden the availability of oncogene-bearing mice and rats to researchers, Taconic will now distribute the v-Ha-ras(TG.AC) OncoMouse under a new licence agreement with DuPont. The company can now distribute the OncoMouse directly to end users under the company's simplified 'conditions-of-use' label licence. This arrangement should make it easier for investigators to acquire access to transgenic models that are proving to be alternatives to traditional animal models used in the two-year carcinogenicity bioassay. Transgenic animal models offer potential benefits that include the use of fewer mice per dosage group and faster screening of compounds. The company supplies murine pathogen free (MPF) laboratory animals including the p53 knockout mouse and the pim-1 transgenic mouse, as well as rats. **Reader Enquiry No. 103**



recombinant samples and in disease states, such as rheumatoid arthritis. High resolution is achieved by exploiting differences in glycan hydrophilicity, which is determined largely by the hydrodynamic volume (size) of the glycan and to a lesser extent by arm specificity and linkage. Hence, glycans with the same hydrodynamic volume, but differing in linkage or anomericity, can be separated. The system can also be used analytically in glycan sequencing either by sequential exoglycosidase digestion or RAAM analysis. The HPLC column should provide an efficient, automated and more detailed means of profiling glycan pools released from monoclonal antibodies, states the manufacturer. Moreover, the column may be used as part of the new multi-dimensional HPLC approach recently developed by the company to separate complex mixtures of oligosaccharides to homogeneity.

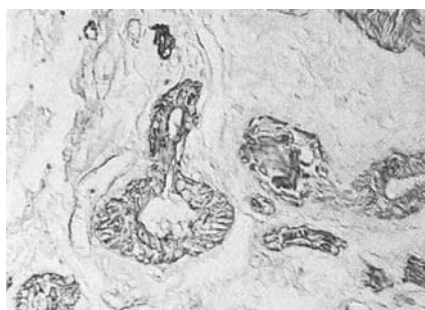
Reader Enquiry No. 104

The **cultivation of hybridoma cells** in the Heraeus miniPerm system is designed to be a simple and economical method for the production of monoclonal antibodies. Typically, as much as 70 mg can be produced in 1-4 weeks, and as the expressed antibody is contained in only 35 ml of medium, downstream processing is said to be straightforward. The company's proSystem media has been optimized for use with the system and, according to the manufacturer, outperforms serum-supplemented medium in terms of antibody production, requiring less than two per cent of the serum needed when using standard medium. The proSystem is available based on DMEM, RPMI or ISCOVES medium and in volumes of 400 ml and 10 litre.

Reader Enquiry No. 105

ELISA kits

KPL's line of HistoMark **immunohistochemical staining systems** offers localization of enzyme activity in frozen and paraffin-embedded tissues, as well as



KPL's HistoMark immunohistochemistry.

other cellular preparations. These are stable liquid reagents for alkaline phosphatase, horseradish peroxidase and β -galactosidase systems. They contain chromogen and activator solutions, buffered substrate solution, counterstain and a detailed protocol (peroxidase systems also contain blocking solution). The manufacturer also offers a range of contrasting colours for double or triple staining. According to the manufacturer, new DAB is stable for one week on the laboratory bench and provides high sensitivity with virtually no background.

Reader Enquiry No. 106

NBS Biologicals has announced the release of seven new **ELISA kits**, further extending their range to include human ICAM-L, VCAM, IL-15, MCP-1 and TNF- α Ultra Sensitive; rat TNF- α Ultra Sensitive; and mouse INF- γ . The kits are all single-plate, 96-test kits for cytokine determination. IL-15 has been shown to have similarity with IL-2, although their *in vivo* roles appear to be different. VCAM and ICAM have been proposed as useful markers for rheumatic diseases, including rheumatoid arthritis. The kits increase the existing selection of human, mouse and monkey products offered by the company to a total of 50.

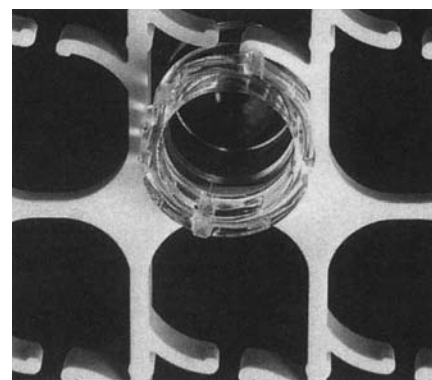
Reader Enquiry No. 107

Two ELISA kits are now available from DAKO for the determination of **total and intact proinsulin**. In normal individuals

proinsulin is only found in small quantities in serum and larger quantities in the β -cells of the pancreas. However, in certain circumstances, such as defects in the processing pathway from proinsulin to insulin, the level of either proinsulin or its processing intermediates may be raised. Both of these assays are based on a flexible, colour-coded microplate-strip format, with a total assay time of three hours. The kits are based on two monoclonal antibodies and are said to show high precision, sensitivity and specificity.

Reader Enquiry No. 108

LockWell immuno-modules from Nalge Nunc are designed for **solid-phase ELISA and RIA applications**. LockWell modules have a standard 96-microwell plate format. Each LockWell frame consists of twelve 1 $\times$ 8-well strips per plate that can be separated into 8 wells, which can be reinserted into the matrix frame. A spring-lock mechanism is designed to ensure that all wells are positioned at the same horizontal level. Modules provide flexibility for solid-phase immunoassays and allow



Nalge Nunc solid-phase ELISA modules.

diagnostic assays from large-scale screening to single-well tests to be performed in one format. LockWell is available in the traditional U-shaped bottom, a C-shaped bottom with a flat centre area and curved edges for efficient washing, and 'star well' with an increased surface area-to-volume ratio.

Reader Enquiry No. 109

Pharmigen has released a new panel of **rat cytokine ELISA reagents**. These include a recombinant protein, a purified capture antibody and a detection antibody for the quantification of rat IL-2, IL-10, TNF- α and MCP-1. Other related rat reagents include IL-10 and GM-CSF, as well as a new PE-conjugated anti-rat GM-CSF.

Reader Enquiry No. 110

Recently, Becton Dickinson Immunocytometry Systems (BDIS) announced the availability of the first reagent system specifically designed for **intracellular cytokine detection in whole blood**. This group of reagents is the newest addition to

ADVERTISEMENTS

NEWCASTLE UNIVERSITY
MOLECULAR BIOLOGY UNIT

PROTEIN SEQUENCING

Membrane-bound, solid or solution samples

£12 per cycle

5 cycles minimum - no set up charge

For further information contact Dr. Joe Gray
Phone: +44 191 222 8612
Fax: +44 191 222 8129
Email: JOE.GRAY@newcastle.ac.uk

READER ENQUIRY NO. 69



Gramsch Laboratories

- Fast Peptide Synthesis: weekend setup:
25 μ moles: 10 AA DM 650.00; 15 AA
DM 730.00; 20 AA DM 780.00

- Antibody Production Service:
Package: DM 2100.00

New Antisera:

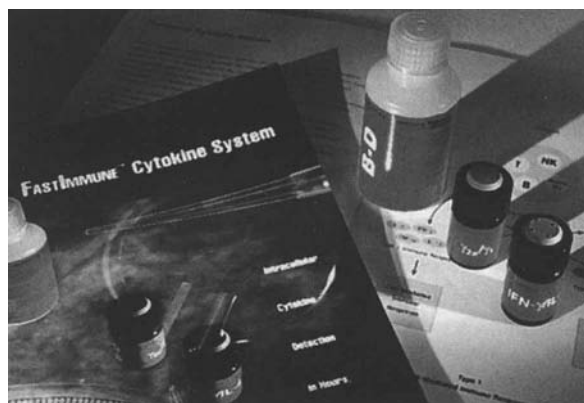
Rab 4/ 5a/ 7 : G-Protein subunit : α 12;
 α 13; α 14; β 2 : γ 2; γ 3; γ 6; γ 7 Nociceptin
Somatost.Rec. mSSTR 1/2A/2B/3

Phone 08138 1769 or 08131 5679-0
Fax 08138 1260 or 08131 5679-24
WWW.ourworld.com/serve.com/
homepages/jgramsch/gi.htm

Gramsch Laboratories

D-85247 SCHWABHAUSEN, Germany

READER ENQUIRY NO. 81



FastImmune helps in the study immune function in normal and pathogenic conditions, such as AIDS and cancer.

the FastImmune family, which includes anti-human IL-2 FITC and PE, anti-human IL-4 PE, anti-human IFN γ FITC and PE, IFN γ FITC/IL-4 PE, g γ ₂ FITC/g γ ₁ PE intracellular control and FACS permeabilizing solution. Designed for type 0, type 1 and type 2 status determinations, these reagents can be useful for studying immune function in normal and pathogenic conditions, such as autoimmune diseases, cancer and AIDS. The system can also help in determining cytokine involvement in disease pathogenesis, T-lymphocyte effector functions, haematopoiesis and thymocyte development. This reagent system measures the expression of cytokines inside cells and provides comprehensive cell-by-cell flow cytometric results from whole blood in only a few hours. Researchers can assess cytokine production by specific cell types using the reagents in this system in combination with cell-surface or intracellular phenotyping.

Reader Enquiry No. 111

CellPath's partnership with Reactifs RAL will permit them to offer a **range of biological stains, reagents and kits** for use in haematology, microbiology, histopathology, cytopathology and parasitology. RAL555 is a rapid alternative to May Grunwald Giemsa (MGG) stain, enabling the differential staining of blood smears, the interpretation of which is identical to classical MGG staining. It can also be used for the staining of gastric samples in the detection of *Helicobacter pylori* before antibiotic treatment. It takes only 3–5 seconds to use each solution in the RAL555 kit, making it suitable for urgent testing and for in-clinic checking of cellularity of fine-needle aspirates. Each RAL555 kit consists of three ready-to-use solutions: alcoholic fixative solution, Ag-buffered eosin solution and Ag-buffered methylene blue solution in 100-ml bottles. The practical design of the kit enables the bottles to be used as staining troughs and the pack can serve as a workstation organizer. All bottles bear CHIP labels for safety and traceability. The solutions are available in

refillable bottles in 100-ml, 1-l and 2-l sizes.

Reader Enquiry No. 112

Monoclonals

Serotec is now able to offer an expanded range of high quality **monoclonal antibodies raised against porcine cell-surface antigens**. The majority of these have been clustered at the 2nd International Swine Cluster of Differentiation Workshop, which has grouped antibodies into clusters (orthologues of human CD antigens), and

also into SWC clusters that recognize antigens that do not directly compare with known human CD antigens. This range of antibodies should allow researchers to identify successfully important leukocyte subsets in the pig, which should further advance the understanding of swine immunology.

Reader Enquiry No. 113

Harlan Sera-Lab now offers Veritype monoclonal antibodies, which are standardized, prediluted, pretested and ready for use in **flow cytometry applications**. These antibodies are provided as unconjugated, purified immunoglobulins, FITC-conjugated reagents and PE-conjugated reagents, which are available in 25- and 100-test sizes. Veritype dual markers (FITC/PE) to human determinants and isotypic negative controls are also available in 50-test sizes. The company can provide Veritype immunological reagents to human, mouse and rat determinants, as well as cell lysing solution.

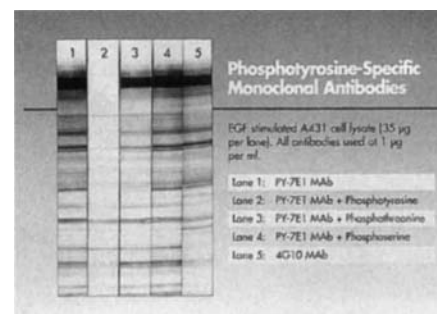
Reader Enquiry No. 114

Kamiya Biomedical has introduced a **Ku monoclonal antibody** for cancer research, which recognizes the 86 K subunit of the

human Ku protein. Ku is a nuclear protein thought to function in cell signalling, DNA replication and transcriptional activation. It serves as the regulatory component of the DNA-associated protein kinase that phosphorylates pol II and transcription factor Sp 1. The clone is raised against a Ku peptide antigen isolated from A431 human epidermoid carcinoma cells and does not react with the Ku 70 K subunit.

Reader Enquiry No. 115

Zymed offers three new epitope affinity-purified, **monoclonal antiphosphotyrosine antibody** products that are effective for studying signal transduction events. The new PY-7E1 clone is stated to have similar reactivity to 4G10 but with a higher sensitivity at lower cost. This antibody reacts with several prominent tyrosine phosphorylated proteins on western blots of A431 cell lysate that can be missed or weakly



Zymed's signal transduction detection.

detected by 4G10. For the broadest antiphosphotyrosine reactivity, the PY-7E1 and PY-1B2 clones are available in a monoclonal cocktail (PY-Plus) that also includes clone PY20. All three reagents are also available as HRP or Sepharose conjugates.

Reader Enquiry No. 116

Antigenix has introduced a selection of **mouse monoclonal antibodies that are paired for immunoassay or ELISA development**. One antibody is supplied uncon-

ADVERTISEMENT

New antibodies for protein trafficking and neuroexocytosis research from AFFINITI

TGN-38, α -actinin, adaptin- α , clathrin, cysteine string proteins, munc18, N-ethylmaleimide-sensitive fusion protein (NSF), SNAP-25, synapsin, synaptobrevin (VAMP), synaptophysin, synaptotagmin, syntaxin, vinculin, *et al.*

For more details on these antibodies and AFFINITI's extensive range of products for neuroscience and cell biology please contact:

AFFINITI Research Products Ltd., Mamhead Castle, EXETER, EX6 8HD, UK
Tel. +44/0 1626 891010; Fax: +44/0 1626 891090

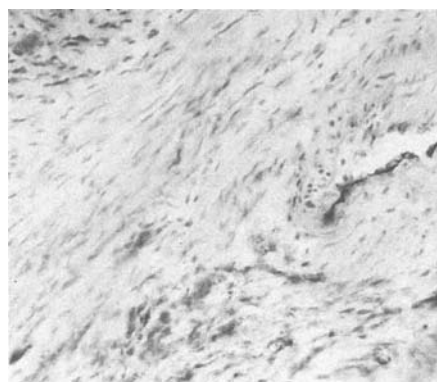
VISIT OUR NEW WEBSITE: <http://www.AFFINITI-RES.COM>

READER ENQUIRY NO. 720

jugated and the other is biotin-labelled for use as a detection reagent. Both of these reagents should prove useful for the development of quantitative ELISAs or immunoassays. Matched antibody pairs are offered, including human VCAM (CD106), serum amyloid P, serum amyloid A, P-selectin (CD62P), ICAM (CD54), β_2 -microglobulin, IL-8, IL-1 β , MCP-1, MCP-3 and EGF. In some cases, the protein standards are also offered.

Reader Enquiry No. 117

Chemicon has introduced six new **monoclonal antibodies to angiotensin-converting enzyme (ACE)** directed against six different epitopes of the ACE molecule. Angiotensin-converting enzyme catalyses both angiotensin I transformation to angiotensin II (a potent vasopressor



Chemicon's ACE antibodies.

agent) and the degradation of bradykinin, a vasodepressor. The monoclonals should be useful for applications such as immunohistochemistry, immunocytochemistry, immunoblotting, immunoprecipitation and ELISAs.

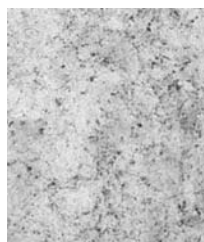
Reader Enquiry No. 118

A new panel of **monoclonal mouse antibodies to human sialosyl-Tn antigen** (NeuAco α 2-6 Gal Nac) is now available from Biomedica for clinical and laboratory researchers. These are IgG-1 isotype monoclonal antibodies specific to human sialosyl-Tn antigen, labelling colon adenocarcinomas and ovarian mucinous carcinomas to produce apical cytoplasmic, membrane or luminal staining patterns. With less frequency, the antibodies are reactive with squamous cell invasive carcinomas. They are also found to be negative with the most of the adult normal human tissues. According to the manufacturer, these monoclonal antibodies to sialosyl-Tn antigen can be used on paraffin-embedded, formalin-fixed and cryostat tissue sections. They are compatible with all of the company's immunohistochemistry staining kits and are available in concentrated, concentrated/dehydrated form, or as a pre-diluted, ready-to-use product.

Reader Enquiry No. 119

Polyclonal antibodies

A **rabbit polyclonal antiserum to glutamic acid decarboxylase (GAD)** is now available from Affiniti Research Products. The antiserum is raised to a synthetic peptide selected from the carboxy-terminus of rat GAD65. This peptide is well conserved in mammals and also exists in the GAD67 molecular form of the enzyme. On western blots of the mammalian central nervous system an intense 65/67 K doublet is obtained at dilutions in excess of 1:10⁵, which is fully blocked in the presence of the immunizing peptide. The antiserum is suitable for use in immunohistochemical applications on frozen, Vibratome and de-waxed tissue sections from a range of species, including rat, feline and human. Intense staining of terminals, fibres and occasionally cell bodies is seen with virtually no background staining at dilutions up to 1:5,000. The antiserum is sold in two sizes (25 μ l and 100 μ l) and the synthetic peptide is also available for use in adsorption (negative) and EIA (positive) applications.



Affiniti's antiserum.

Reader Enquiry No. 120

Incstar has announced the release of an affinity-purified, **rabbit polyclonal antibody to 5HT₂ receptor** for immunohistochemistry. The 5HT₂ receptor is present in the brain and spinal cord, the enteric nervous system of the gastrointestinal tract, the vasculature of central and peripheral blood vessels, as well as other organ systems. The distribution of 5HT₂ receptor has been studied extensively by receptor binding and *in situ* hybridization. High concentrations of this receptor are found in the cortex, the amygdala, the hippocampus and a number of cranial nuclei and reticular formations. The availability of a histochemical antibody to 5HT₂ receptor is stated to provide for higher resolution than other localization methods. The antiserum to 5HT₂ receptor was generated in rabbit against a multiple antigenic peptide of an amino-terminal synthetic sequence from rat.

Reader Enquiry No. 121

Lifescan has an agreement to distribute the Alpha 4 and Alpha 4 LS microplate-based **immunoassay instruments** for SFRI Laboratories in the UK. These instruments are fully automated microplate processing systems, utilizing user-friendly software and flexible hardware to make the systems better suited to

microplate-based immunoassays. The company offers protocols for its entire range of products.

Reader Enquiry No. 122

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

SYNTHETIC VIRAL PROTEINS

HIV-1 TAT HB10

(1000 USD/100 μ g)

HBV pre-S1 adw

(1000 USD/100 μ g)

myristil-HBV pre-S1 adw

(800 USD/50 μ g)

AND RELATED ANTIBODIES

Fax 39 823 612222

Email info@tecnogen.it

TECNOGEN SCpA - Science Park

81015 Piana di M. Verna - CE - Italy

WEB Site: <http://www.tecnogen.it>

READER ENQUIRY NO. 80

CUSTOM PEPTIDE SYNTHESIS

We take good care of your product

QUANTUM
BIOTECHNOLOGIES INC.

Choose from 4 levels
of purity

Prices as low as
\$18 per residue
for 20-30 mg

Call for our flyer, toll free:

1-888-DNA-KITS

Fax: 888-688-3785

E-mail: info@qbi.com

Peptides are produced with Fmoc chemistry.
Numerous peptide modifications available.

READER ENQUIRY NO. 651